

Figure 1

BMY_HPP1_A

```

1 CTAGTTTACT TCTACAATT CGGATGGAAG GATTATGGTG TAGCGTCCT TACTACTATC 60
1 L V Y F Y N F G W K D Y G V A S L T T I 20

61 CTAGATATGG TGAAGGTGAT GACATTGCCC TTACAGGAAG GAAAAGTAGC TATCCATTGT 120
21 L D M V K V M T F A L Q E G K V A I H C 40

121 CATGCAGGGC TTGGTCGAAC AGGT 144
41 H A G L G R T G 48

```

BMY_HPP1_B

```

1 GATGTCCTCT GGGCCCTCCT GTGGAACACA GTT 33
1 D V F W A L L W N T V 11

```

Figure 2

1 GTGGCCCGGAGGCGCCGAGGCCAGGTAGGTGCGATGGGCGTGCAGCCCCCAACTTCTC 60
 1 W P G R R R R G Q V G A M G V Q P P N F S 20

 61 CTGGGTGCTTCCGGGCCGGCTGGCGGACTGGCGCTGCCGCGGCTCCCGCCCCACTACCA 120
 21 W V L P G R L A G L A L P R L P A H Y Q 40

 121 GTTCTGTGTGGACCTGGGCGTGGCGCACCTGGTGTCCCTGACGGAGCGGGGCCCCCTCA 180
 41 F L L D L G V R H L V S L T E R G P P H 60

 181 CAGCGACAGCTGCCCCGGCCTCACCCCTGCACCGCCTGCGCATCCCCGACTTCTGCCCGCC 240
 61 S D S C P G L T L H R L R I P D F C P P 80

 241 GGCCCCCGACCATCGACCCGCTTCGTGCAGATCGTGGACGAGGCCAACGCAGGGGAGA 300
 81 A P D Q I D R F V Q I V D E A N A R G E 100

 301 GCGTGTGGAGTGCACCTGTGCTCTGGGCTTTGGCCGCACTGGCACCATGCTGGCCTGTTA 360
 101 A V G V H C A L G F G R T G T M L A C Y 120

 361 CCTGGTGAAGAGCGGGGCTTGGCTGCAGGAGATGCCATTGCTGAAATCCGACGACTACG 420
 121 L V K E R G L A A G D A I A E I R R L R 140

 421 ACCCGGCCCATCGAGACCTATGAGCAGGAGAAAGCAGTCTTCCAGTCTTACCAGCGAAC 480
 141 P G P I E T Y E Q E K A V F Q F Y Q R T 160

 481 GAAATAAGGGGCTTAGTACCCCTTCTACCAGGCCCTCACTCCCTTCCCCATGTTGTGCA 540
 161 K * G A L V P F Y Q A L T P L P H V V D 180

 541 TGGGGCCAGAGATGAAGGGAAGTGGACTAAAGTATTAAACCCCTCTAGCTCCCATTTGGCTG 600
 181 G A R D E G K W T K V L N P L A P I G * 200

 601 AAGACACTGAAGTAGCCACCCCTGCAGGCAGGTCTGATTGAAGGGGAGGCTTGTA CTG 660
 201 R H * S S P P L Q A G P D * R G G L Y C 220

 661 CTTTGTGAATAAATGAGTTTACGAACCAAGGAAAAAAGAAAAA 720
 221 F V E * M S F T N Q G K K K K K K R K K 240

 721 AAAAAAAAAAAAAAAAAAAGAA 746
 241 K K K K K K K R 248

Figure 3

```

1 ATGGCTAGAA TGAACCTCCC TGCTTCTGTG GACATTGCAT ACAAAAAATGT GAGATTTCCT 60
1 M A R M N L P A S V D I A Y K N V R F L 20

61 ATTACACACA ACCCAACCAA TACCTACTTT AATAGATICT TACAGGAACT TAAGCAGGAT 120
21 I T H N P T N T Y F N R F L Q E L K Q D 40

121 GGAGTTACCA CCATAGTAAG AGTATGAAAA GCAACTTACA ACATTGCTCT TTTAGAGAAG 180
41 G V T T I V R V * K A T Y N I A L L E K 60

181 GGAAGCATCC AGGTTCCGGA CTGGCCTTTT GATGATGGTA CAGCACCATC CAGCCAGATA 240
61 G S I Q V P D W P F D D G T A P S S Q I 80

241 ATTGATAACT GGTAAAACT TATGAAAAAT AAATTCATG AAGTCCCTGG TTGTTGATT 300
81 I D N W L K L M K N K F H E D P G C C I 100

301 GCAATTCACT GTGTGTAGG TTTTGGGTGA GCTCCAGTTG CTAGTTGCCC TAGCTTTAAT 360
101 A I H C V V G F G * A P V A S C P S F N 120

361 TGAAGGTGGA ATGAAATATG AAAATGTAGT ACAGTTCATC AGATAAAAGT GACATGGAAC 420
121 * R W N E I * K C S T V H Q I K V T W N 140

421 TTTTAACAGC AAACAACCTT TGTATTGGA GAAATATTGT CTTAAAAATAT GCTTGACCT 480
141 F * Q Q T T F V F G E I L S * N M L A P 160

481 CAGAAATCCC AGAAATAACT GTTTCCTTCA G 511
161 Q K S Q K * L F P S 171

```

Figure 4A

1	CTCAGGCAGA	ACTGTAGGCG	CAAGATGTGCT	CATGCGCACCC	AGGCTTTCTT	TTTGAATTC	60
1	L R Q N	Y E A	K S A	H A H Q	A F F F	L K F	20
61	GAGGAGCTGA	AGGAGGTGAG	CAAGGACAG	CCGAGACTGG	AGGCTGAGTA	CCCTGCCAAC	120
21	E E L K	E V S	K E Q	P R L E	A E Y	P A N	40
121	ACCACCAAGA	ACTGTTAAAC	ACATGTGCTA	CCCTATGACC	ACTCTCAGGGT	CAGGCTGACC	180
41	T T K N	C * P	H V L	P Y D H	S R V	R L T	60
181	CAGCTGGAGG	GAGAGCCCTA	TTCTGACTAC	ATCAATGCCA	ACTTGTGCC	AGGCTACACC	240
61	Q L E G	E P H	S D Y	I N A N	L V P	G Y T	80
241	CGCCACAGG	AGTTCTATGC	CTCTCAGGG	CCTCTCAAGA	AAACACTGGA	GAACCTCTGG	300
81	R P Q E	F I A	S C Q	P L K K	T L E	N F W	100
301	CGGCTGGTGC	GGAGCAGCA	GGTCCGCATC	ATCATCATGC	CGACCATCAG	CATGGAGAAC	360
101	R L V R	E G Q Q	V R I	I I M P	T I S	M E N	120
361	GGGAGGGTGC	TGTGTGAGCA	TTACTGGCTG	ACCGACTCTA	CCCCGGACAC	CCATGGTCAC	420
121	G R V L	C E H T	Y W L	T D S T	P D T	H G H	140
421	ATCAOACTCC	AOCTCTAGTC	TGAGGAGCCT	GAGGATGAGT	GGACCAAGCG	GGAATTCGAG	480
141	I T I H	L L A	E E F	E D E W	T K R	E F Q	160
481	CTGCAGCAGC	TTGTQACGA	ACATCAACGG	AGGGTGGAGC	AACTCGAGTT	CACCAACGAT	540
161	L Q H V	V T C K	H Q R	R V E Q	L G F T	T T T *	180
541	TCCGACCACA	GCACTCTTGA	GGCTCCAGAC	TCCCTGCTCG	CCCTTATGGA	GCTGGTACAG	600
181	S D H S	I L E	A P S	S L L A	C F M E	L V Q	200
601	TAGCAGGCAG	GGGCCACCCA	GGGCGTGGGA	CCCATCTCGG	TGCATCTGAG	GGGCTGTCCC	660
201	* Q A R	A T Q	G V G	P I L V	H C R	G C P	220
661	TGCGGTGTGG	GCATGGGCGC	GACAGGCACC	TTCTGTGCCCC	TGTCTGAGGT	GCTGCAGCAG	720
221	C G V G	M G R	T G T	F V A L	S R L L	L Q Q	240
721	CTGGAGGAGG	AGCAGATGGT	AGACGTGTGT	CATGCTGTGT	ATGCAGCTCG	GATGCAOACG	780
241	L E E E	Q H V	D V F	H A V Y	A L R	M H Q	260
781	CCCTCATGAA	TCCAGACCTT	GAGCCAGTAG	GTCTTCTGTC	ACAGCTGCCT	ACTGAACAG	840
261	P L M I	Q T L	S Q Y	V F L H	S C L	L N K	280
841	ATTCTGGAAG	GACCTCTCAA	CATCTCTGAG	TCITGGGCCA	TCITCTGTAC	GGACCTCCCG	900
281	I L E G	F F N	I S E	S N P I	S V T	D L P	300
901	CAGGCGTGTG	CCAAAGGKCA	AGCCACTGCC	AAATGCTGCT	TCCTTGAAGA	GTACGAGGCC	960
301	Q A C A	K R A	A S A	N A G F	L A E	Y E A	320
961	ATCAAGGAGC	AGGCTGGCTT	TTCCGACACC	CCGCTGTGCT	ATGAGCAGGA	CAGCCCGCGT	1020
321	I K D E	A G F	S A P	P P G Y	E Q D	S P V	340
1021	TCCTATGACC	GTTCTCAGGG	GCAGTTTTCT	CCGGTGGAGG	AGAGCCCCCC	TGACGACATG	1080
341	S Y D R	S Q Q	Q F S	P V E E	S S P P	D D M	360
1081	CCTCTCTGGA	AGCCAATGAT	CTGTGCTCTG	CAGGGGTGGC	CCTCTGPGCC	TGATCATAGC	1140
361	P L W K	P M I	C A L	Q G G P	S G R	D H T	380

Figure 4B

```

1141 GTGCTGACTG GCCCGCAGG GCCAAGGAG CTCTGGGAGC TGGTGTGGCA GCACAGGGCT 1200
381 V L T G P A G P K E L W E L V W Q H R A 400

1201 CATGTGCTTG TCTCTCTTTG CCCACCCAAT GTCATGGAGA AGGAATTCTG GCCAACGGAG 1260
401 H V L V S L C P P N V M E K E F W P T E 420

1261 ATGCAGCCCG TAGTCACAGA CATGGTGACG GTGCACTGGG TGGCTGAGAG CAGCACAGCA 1320
421 M Q P V V T D M V T V H W V A E S S T A 440

1321 GGCTGSETCT GTACCCTCCT CAGGGTCACA CATGGGGAGA GCAGGAAGGA AAGGGAGGTG 1380
441 G W F C T L L R V T H G E S R K E R E V 460

1381 CAGAGACTGC AATTTCATA CCTGAGGCCT GGGCATGAGC TGCOCGCCAC CACCTGCTG 1440
461 Q R L Q F P Y L E F G H E L P A T T L L 480

1441 CCCTTCCTGG CTGCTGTGGG CCAGTGCTGC TCTCGGGGCA ACACAAAGAA GCCGGGCACA 1500
481 P F L A A V G Q C C S R G N N K K P G T 500

1501 CTGCTCAGCC ACTCCARCA GGGTGCAACC CAGCTGGGGA CCTTCCTGSC CATGGAGCAG 1560
501 L L S H S N K G A T Q L G T F L A M E Q 520

1561 CTGCTCAGCG AGGCAGGTC TGAATGCACC GTGGATATCT TTAACCTGGC CCTGCAGCAG 1620
521 L L Q Q A G S E C T V D I F N V A L Q Q 540

1621 TCTGAGCCT GTGGCCTTAT GACCCCAACA CTGAAGCAGT ATGTCTACCT CTACAACCTG 1680
541 S Q A C G L M T P T L K Q Y V Y L Y N C 560

1681 CTGAACAGCG CGCTGGCAGA CGGGCTGCC 1710
561 L N S A L A D G L P 570

```

1	ATGTTCACTTTTAAAAAACTTCAGGATGGGCACAAACACACAGAAGTGGGAAATGAATAAA	60	
61	AGAGTATTGATAAATTTTGGAAAATGTTGAAGCTGAGTAATGGGCTTTTCAGTCCAGTGT	120	
121	AAAGCTGTGGAGCGCGGGAGCAAAGGTAAGAATGATGTAATGCGCTGGCTGTCCAAA	180	
181	GCATCTTTTGTGTGGAATGGTTATTCCAGTCATCTCTTTATGAATCAAATGTGAGGGGC	240	
241	TGCTTTGTGGACGGAGTCCTTTGCAAGAGCAGATCAACGGGAAAGAGAAAGAGACATTCA	300	
301	CTTGGAGGGCTCTTGCTGAAAATGGGTTTAACTCTCTCTTTTGCCAGTCACCAACAGCCTG	360	
361	ACCTCATACACTTTTAGTACAATGGAGTGGCTGAGCCTTTGAGCACACCACCATTACATC	420	
421	ATCGTGGCAAATTAAAGAAGGAGGTGGGAAAAGAGGACTTATTGTTGTGATGGCCCCATGA	480	
1		M A H E	4
481	GATGATTGGAACCTCAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAAAGTGGAAACGGA	540	
5		M I G T Q I V T E R L V A L L E S G T E	24
541	AAAAGTGTGCTAATTGATAGCCGGCCATTTTGTGGAATACAATACATCCACATTTTGGGA	600	
25		K V L L I D S R P F V E Y N T S H I L E	44
601	AGCATTAAATATCAACTGCTCCAAGCTTATGAAGCGAAGGTTGCAACAGGACAAAGTGTT	660	
45		A I N I N C S K L M K R R L Q Q Q D K V L	64
661	AATTACAGAGCTCATCCAGCATTGAGCGAAACATAAGGTTGACATTGATTGCAGTCAGAA	720	
65		I T E L I Q H S A K H K V D I D C S Q K	84
721	GGTTGATGTTTACGATCAAAGCTCCCAAGATGTTGGCTCTCTCTCTTCAGACTGTTTCT	780	
85		V V V Y D Q S S Q D V A S L S S D C F L	104
781	CACTGTACTTCTGGGTAACCTGAGAGAAGGCTTCAACTCTGTTACCTGCTTGCAGGTGG	840	
105		T V L L L G K L E K S F N S V H L L A G G	124
841	GTTTGCTGAGTTCTCTCGTTGTTCCCTGGCCTCTGTGAAGGAAAAATCCACTCTAGTCCC	900	
125		F A E F S R C F P G L C E G K S C T L V P	144
901	TACCTGCATTTCTCAGCCCTTGCTTACCTGTTGCCAACATTGGGKCAACCCGAATCTCTCC	960	
145		T C I S Q P C L P V A N I G P T R I L P	164

Figure 5B

961 CAATCTTTATCTTGGCTGCCAGCGAGATGTCCTCAACAAGGAGCTGATGCAGCAGAATGG 1020
 165 N L Y L G C Q R D V L N K E L M Q Q N G 184

1021 GATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATCCCCGAGTC 1080
 185 I G Y V L N A S N T C P K P D F I P E S 204

1081 TCATTTCCCTGCGTGTCCCTGTGAATGACAGCTTTTGTGAGAAAAATTTGCCGTGGTTGGA 1140
 205 H F L R V P V N D S F C E K I L P W L D 224

1141 CAAATCAGTAGATTCATTGAGAAAGCAAAGCCTCCAATGGATGTGTTCTAGTGCACTG 1200
 225 K S V D F I E K A K A S N G C V L V H C 244

1201 TTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCTACATCATGAAGAGGATGGA 1260
 245 L A G I S R S A T I A I A Y I M K R M D 264

1261 CATGCTCTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACTATATCTCCAAA 1320
 265 M S L D E A Y R F V K E K R P T I S P N 284

1321 CTTCAATTTTCTGGGCCAACTCCTGGCCTATGAGAAGAAGATTAAGAACCAGACTGGAGC 1380
 285 F N F L G Q L L A Y E K K I K N Q T G A 304

1381 ATCAGGGGCCAAAGAGCAAACCTCAAGCTGCTGCCCTGGAGAGCCAAATGAACCTGTCCC 1440
 305 S G P K S K L K L L P L E K P N E P V P 324

1441 TGCTGTCTCAGAGGGGTGGACAGAAAAGCGAGACGCCCTCAGTCCACCTGTGCCGACTC 1500
 325 A V S E G G Q K S E T P L S P P C A D S 344

1501 TGCTACCTCAGAGGCGCAGGACAAAGGCCCGTGCATCCCCGCCAGCGTGCCAGCGTGCC 1560
 345 A T S E A A G Q R P V H P A S V P S V P 364

1561 CAGCGTCAGCGCTCGCTGTAGAGGACAGCCCGCTGGTACAGGCGCTCAGTGGGCTGCA 1620
 365 S V Q P S L L E D S P L V Q A L S G L H 384

1621 CCTGTCCGCAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTTCCTCTCTCTGGATAT 1680
 385 L S A D R L E D S N K L K R S F S L D I 404

1681 CAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTACATGGCTTCTCTCATCAGA 1740
 405 K S V S Y S A S M A A S L H G F S S S E 424

1741 AGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGATGGGACCAACAAGCTATGCCA 1800
 425 D A L E Y Y K P S T T L D G T N K L C Q 444

1801 GTTCTCCCTGTTCAGGAACATATCGGAGCAGACTCCCGAAACCAAGCTCCTGATAAGGAGGA 1860
 445 F S P V Q E L S E Q T P E T S P D K E E 464

Figure 5C

1861	AGCCAGCATCCCCAAGAAGCTGCAGACCGCCAGGCCTTCAGACAGCCAGAGCAAGCGATT	1920
465	A S I P K K L Q T A R P S D S Q S K R L	484
1921	GCATTCGGTCAGAACAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTATCTCCACTGCA	1980
485	H S V R T S S S G T A Q R S L L S P L H	504
1981	TCGAAGTGGGAGCGTGGAGGACAATTACCACACCAGCTTCCTTTTCGGCCTTTCCACCAG	2040
505	R S G S V E D N Y H T S F L F G L S T S	524
2041	CCAGCAGCACCTCACGAAGTCTGCTGGCCTGGGCCTTAAGGGCTGGCACTCGGATATCTT	2100
525	Q Q H L T K S A G L G L K G W H S D I L	544
2101	GGCCCCCAGACCTCTACCCCTTCCTGACCAGCAGCTGGTATTTTGGCCAGAGTCTCTC	2160
545	A P Q T S T P S L T S S W Y F A T E S S	564
2161	ACACTTCTACTCTGCCTCAGCCATCTACGGAGGCAGTGCAGTTACTCTGCCTACAGCTG	2220
565	H F Y S A S A I Y G G S A S Y S A Y S C	584
2221	CAGCCAGCTGCCCACTTCCGGAGACCAAGTCTATTCTGTGCGCAGGCGGCAGAACGCAAG	2280
585	S Q L P T C G D Q V Y S V R R R Q K P S	604
2281	TGACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTTGAAGAGCAGTTTAA	2340
605	D R A D S R R S W H E E S P F E K Q F K	624
2341	ACGCAGAAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTACAGAGAACAGGTCACGGGA	2400
625	R R S C Q M E F G E S I M S E N R S R E	644
2401	AGAGCTGGGGAAAGTGGGCAGTCAGTCTAGCTTTTCGGGCAGCATGGAATCATTTAGAGGT	2460
645	E L G K V G S Q S S F S G S M E I I E V	664
2461	CTCCTGAGAGAAAGACACTTGTGACTTCTATAGACAATTTTTTTTTCTGTTCACAAA	2520
665	S	665
2521	AAATTCCCTGGGAATCTGAAATATGTATGTGGGCATACATATATATTTTGGAAAATGGA	2580
2581	GCTATGGTGTAAAAGCAACAGGTGGATCAACCCAGTTGTTACTCTCTTAACATCTGCATT	2640
2641	TCAGAGATCAGCTAATACTTGCTCTCAACAAAAATGGAAGGCAGATGCTAGAATCCCCC	2700
2701	CTAGACGGAGGAAAACCATTTTATTTCAGTGAATTACACATCCTCTTGTCTTAAAAAAGC	2760
2761	AAGTGTCTTTGGTGTGGAGGACAAAATCCCTACCATTTTCACGTTGTGCTACTAAGAG	2820

Figure 5D

2821	ATCTCAAAATATTAGTCTTTTGTCCGGACCCCTTCCATAGTACACCTTAGCGCTGAGACTGAG	2880
2881	CCAGCTTGGGGGTGAGGTAGGTAGACCTGTTAGGGACAGAGCCTAGTGGTAAATCCAAG	2940
2941	AGAAATGATCCTATCCAAAGCTGATTCACAAACCCACGCTCACCTGACAGCCGAGGGACA	3000
3001	CGAGCATCACTCTGCTGGACGGACCATTAGGGGCCCTTGCCAAGGTCTACCTTAGAGCAAA	3060
3061	CCCAGTACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCATATCTGCTAGCCCATTT	3120
3121	TCTAGGCATTGTGAATAGGTAGGTAGCTAGTCACACTTTTCAGACCAATTCAAACGTGCT	3180
3181	ATGCACAAAATTCGCCGTGGGCTAGATGGAGATAATTTTTTTTCTTCTCAGCTTTATGA	3240
3241	AGAGAAGGGAAACTGTCTAGGATTCAAGCTGAACCAACAGGAACCTGGCAACATCACGATT	3300
3301	TAAAGCTAAGGTTGGGAGGCTAACGAGTCTACCTCCCTCTTTGTAAATCAAAGAATTGTTT	3360
3361	AAAATGGGATTGTCAATCCTTTAAATAAAGATGAACCTGGTTTCAAGCCAAATGTGAATT	3420
3421	TATTTGGGTTGGTAGCAGAGCAGCAGCACCTTCAAATTTCTCAGCCAAAGCAGATGTTTTT	3480
3481	GCCCTTCTGCTTCACTGCATGGGATACAGTTGGTAAAAATGTAATAATATGGCAGAATTTT	3540
3541	ATAGGAAACTTCCTAGGGAGGTAAATTATGGGAAGATTAAAGAAAGGTACAAATTGTCTGAG	3600
3601	GAGAAGCAGGAAACCTGTTTCTTCTAGTGGCTTTTATCCCTCGGCATGCGATGGGGCTGA	3660
3661	TGTTTCTATATTAATGGCTCAGACTTTCACATTTACTAGTAGGGCTGAGAGAGGCTTTTAGTG	3720
3721	AGGAAAGAATATTCAGAATAAAACGGTTGAGAAAGCTGAGAAGACCATTGAGTTTGTATC	3780
3781	AGTTGTGAATAGAGTGCAAGCCATGCCAAGCTGTTTTTGGAAACGCTGGCCGGCGTGT	3840
3841	CTTCAGTGGAAAAAGCAAATCAAATGGAGCGAGAGCAAAGGGGCGTCCCTCAGTCTCTCAA	3900
3901	CCTACAATCACTGTATGGAATCGGTCTCTGGCAGCTGAACATAGGAGGTCACTGGAACAAAG	3960
3961	TGATAGTGCAGATTGGCTTTCAAACATCTCTGGCTTGAGTTTTATCAGCTACAATGTG	4020

Figure 5E

4021 GGTCTCTCTTTTGAAGCCTTAATTCAACAACAGCAGCTTTTGGGGGTGGGGCTGGGCGGGT 4080
 4081 GTTGTCAATTGTTCTTTCCCTTCCTGTAGTGTCTGCTAGTTGCTGCCTCGTATCTCAGGTT 4140
 4141 TTTCTCTGTTTTTGAGAAATGGACAGTTTTTTGACCAGGATGTGACTTCATGTTTCCTAT 4200
 4201 GGTGACTTCTAAAACAGCACAGAATGATATGACTCAACACAGACCGACTTGGTTATGGG 4260
 4261 GATGATGAGCCGCACAGACCTCACTAGTTGTGCACAAATAATGTCTATGATGGGGTGA 4320
 4321 AAGTGAAGGCAGAAGAGGGTCAGCCGCAATTGTTATGATATGGGAAAGTGCCGGTCAACG 4380
 4381 ATTTGAGTTAGTTTTTAGATATACATTGAATCTTTAATCAGACATTCTCAAGTTTCACA 4440
 4441 CAGTAGTTTTTGATGTTATGTACACACACACCAAAATGTGTAAACAGTTCAGCACTTCCAGA 4500
 4501 GTGTGGTCATGCCCAAAACATGTTTAAAGAAAGAAAGCAGTAGCTCCTTGCTAACGATGT 4560
 4561 TTCAGGAGGTTTGGGGCACTTGGTTTTAATGAGCTTCTGTCAATTTAGGGTTCTCTTGGC 4620
 4621 CATGGTCCCTTCCTTCTGGAACGTGTGATGTAGTCACATCCTACAGCCTTGTAGTCTGGT 4680
 4681 TCACTAGTGTGAGATAATCAGTTCTTGGAAATCGAGACTGCCGTGGCGAAGGGTGGCCTC 4740
 4741 GGAGGCAGGCTCTGGAGCTGCTTGGATGTCTTTAGGTGGGGTGGTGGCTGGCTCTCTTCA 4800
 4801 GCATGTAATTGGGGAACCTCGCGCTACTAGGGGTGATACAGATGGTGATTTTAAAGA 4860
 4861 GCAAACTAGACTTCTATGTGAGAAGTGCTGGAAAATGATTTAGGACGTGTAAAGTTAGA 4920
 4921 TGGAAAGACTGTAATGTTTAATATGAATATAGTGTTCTTTTGAAGTAAGGCCAGCTGTT 4980
 4981 GAACGGTTAACTGTGCATTCTCATTGTTGATGTGCATGTATGTTAATGTATGAAATGA 5040
 5041 TTAATAAAATCAAACTGGTACCTGTTTATCCATAAAAAAAAAAAAAAAAAAAAAAAAAA 5100
 5101 AAAAAAAAAAG 5111

Figure 6A

BMV_HPP1_FL	(1)	1	50
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(1)	MGHLPTGIHGARRLLFLLLWFLVLFKNATAFHVTVQDDNNIVVSLEASDVIT	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----
BMV_HPP1_FL	(1)	51	100
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(51)	SPASVYVVVKITGESKNYFFEFEEFNSTLPPPVIFKASYHGLYYIITLVVV	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----
BMV_HPP1_FL	(1)	101	150
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(101)	NGNVVTKPSRSITVLTLPFLVTSVSIYDYKPSPETGVLFIEHYPEKYNVF	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----
BMV_HPP1_FL	(1)	151	200
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(151)	TRVNISYWEKGDFRTMLYKDFFGKKTIVENHWLPGMCYSNITQLVSEATF	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----
BMV_HPP1_FL	(1)	201	250
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(201)	NKSTLVEYSGVSHEFKQHRTAPYPQNISVRIVNLKNNWEEQSGNFEE	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----
BMV_HPP1_FL	(1)	251	300
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(251)	SFMRSQDTIGKEKLFHFTETPEIPSGNISSGWPDNSSDYETTSSQPYWW	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----
BMV_HPP1_FL	(1)	301	350
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(301)	DSASAAPESSEDEFVSVLPMEYENNSTLSETEKSTSGSFSFFFPQVQMILTWL	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----
BMV_HPP1_FL	(1)	351	400
BMV_HPP1_A	(1)	-----	-----
BMV_HPP1_B	(1)	-----	-----
HS_RPTPO	(351)	PPKPPTAFDGEHIIHIEREENFTEYLMVDEEAHEFVAELKEPGKYKLSVTT	-----
MM_RPTPO	(1)	-----	-----
PYP3_SP	(1)	-----	-----

401

450

451

500

501

550

551

600

601

650

651

700

701

750

FIGURE 6C

		801		850
BMV_HPP1_FL	(137)	KILLDAENRPTMKDVSEGPSLELEKTMSEMTVQKEPLAHDSVSN		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(1)	-----		
HS_RTPPO	(801)	PSVPTTFLAVSTVETEMNNVVVSVVAILSLTGLGLVLTILLRKKHLS		
MM_RTPPO	(1)	-----		
PYP3_SP	(1)	-----		
		851		900
BMV_HPP1_FL	(187)	NPMTVAIDIDNPGMSNQQCFDPIQKRRN---VCLQPLTLHKRLRYS		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(1)	-----		
HS_RTPPO	(851)	MARECCGTTVNFASERERCKLPYNRS		
MM_RTPPO	(40)	MARECCGTTVNFASERERCKLPYSRRSVFALLTILPSCIAITVLLAFY		
PYP3_SP	(1)	-----		
		901		950
BMV_HPP1_FL	(233)	SULKRAENLEQGEQTVTQAQILVGHKPKROKLSCYIQPSPPDHKEFL		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(1)	-----		
HS_RTPPO	(978)	-----KNGIKKKRIKTHIQOLDPDAIKDKMAKDSYKFSQOFELKIKT		
MM_RTPPO	(90)	INPWSKNGCKKKRIKTHIQOLDPDSYIKDKMAKDSYKFSQOFELKIKT		
PYP3_SP	(1)	-----HSFKEVSTNGVLTPLTIKKAYMILEGNEEETPL		
		951		1000
BMV_HPP1_FL	(283)	VRTSLFWSQKFGGLGLKNGSPINHSIPKEAGQSGFADYSSSISSEG		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(1)	-----		
HS_RTPPO	(923)	DIPHFADPLPNRCINRYNINLPYFSEVSLVSNNEEESADYNANNVTE		
MM_RTPPO	(140)	DIPHFADPLPNRCINRYNINLPYFSEVSLVSNNEEESADYNANNVTE		
PYP3_SP	(38)	LNTRLKLSKKALANRYNINLPYENHVRPLDQKWK-SADYINASTLCKI		
		1001		1050
BMV_HPP1_FL	(333)	EPVSPFANHHPEINSAHQVHCCKTKEVGSVSGVQNSRTPRSLDQGS		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(1)	-----HVFVAWALNNTV		
HS_RTPPO	(973)	YNSPQELTATQGHLEIRNDEWKMVLQCK-SLTIVMLTQCNKLRKCKCD		
MM_RTPPO	(190)	YNSPQELTATQGHLEIRNDEWKMVLQCK-SLTIVMLTQCNKLRKCKCD		
PYP3_SP	(87)	P-SCKTETATQGHTSNELIVFWKMYVSPKESGLIVMLTKLRKRRKCKD		
		1051		1100
BMV_HPP1_FL	(383)	KQAFLVFHEETQDSKDSAASSHQSELSAARRLLAAKALANINVEKEE		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(12)	-----		
HS_RTPPO	(1021)	HFWFFPEPIAYGDITVEKESSE-----EQD-IWACERSR		
MM_RTPPO	(238)	HFWFFPEPIAYGDITVEKESSE-----EEE-IWASERSR		
PYP3_SP	(136)	HTWVELFTFLNIGDSLVDVYKVT-----LTSLNIVQVVESE		
		1101		1150
BMV_HPP1_FL	(433)	LRKRVKQWQKINSRDGPRICRIGERPFLICLSMMSVEXLEPVITKEVD		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(12)	-----		
HS_RTPPO	(1055)	LYALMDMDHSLNTPDPSSTPTANSAFSTQPHHMQROATKSKKSL		
MM_RTPPO	(272)	LYALMDMDHSLNTPDPSSTPTANSAFSTQPHHMQROATKSKKSL		
PYP3_SP	(174)	SKDGKVKKLLHVNINQDTEHTFTSLSTKTKSKLSYSPDFETSR		
		1151		1200
BMV_HPP1_FL	(483)	HTVDPDADALFPLTEKGQHTLICVLITQINQTVQVEEFLAHAEAF		
BMV_HPP1_A	(49)	-----		
BMV_HPP1_B	(12)	-----		
HS_RTPPO	(1105)	THCSACGRGTGTFALDRITQHIRDFPDLGLVSENRRYRMSMTE		
MM_RTPPO	(322)	THCSACGRGTGTFALDRITQHIRDFPDLGLVSENRRYRMSMTE		
PYP3_SP	(224)	THCSACGRGTGTFALDRITQHIRDFPDLGLVSENRRYRMSMTE		

FIGURE 6D

		1201		1237
EMY_HPP1_FL	(533)	KVNFDSENGPTLVNTHQIFKHTSEKRRMTDGP	KPGL	
EMY_HPP1_A	(49)	-----		
EMY_HPP1_B	(12)	-----		
HS_RPTPO	(1155)	EQYIFHEDGCKMWLSEKQQFCISDIYENVSKS---		
MM_RPTPO	(372)	EQYIFHEDGCKMWLSEKQQFCISDIYENVSKS---		
PYP3_SP	(274)	QSVQSWDGLVETVSGELLQKEFMPQL-----		

1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300

Figure 7A

```

      1                                     50
BMY_HPP2_FL (1) -----
BMY_HPP2.partial (1) -----
HS_CDC14A (1) -----
HS_CDC14B (1) MKRKSERRSSWAAAPCSRRCSSSTSPGVKKLRSTQQDPRRRDQDDVYL
SC_CDC14 (1) -----MRRSVLYDNTIE

      51                                     100
BMY_HPP2_FL (1) -----
BMY_HPP2.partial (1) -----
HS_CDC14A (1) MKDGLYFTLLNRKSTVNTHYSLDEEYVENYADFGPNTLMVYK
(51) DDTDSMCFRILYSRKASASNVBYESTINELEYENGVDFGPNTLMVYK
HS_CDC14B (13) FMRGRVYLHAYDYTIEDTDELVFTVYDALIVYSGHLDFGPNTLMVYK
SC_CDC14 (13) -----

      101                                    150
BMY_HPP2_FL (1) -----
BMY_HPP2.partial (1) -----
HS_CDC14A (50) CCKLNKKKSYSLSRKKLVHNTCFDCKRANARFLGAVAYTLKPTDEE
HS_CDC14B (101) CCKLNKKKSYSLSRKKLVHNTCFDCKRANARFLGAVAYTLKPTDEE
SC_CDC14 (63) AVIFHEIINDPENANKAVVYFSSASTQRANARCNLCVAVLVQAWVHQ

      151                                    200
BMY_HPP2_FL (1) -----
BMY_HPP2.partial (1) -----
HS_CDC14A (100) AYRALSGSNPPYLPFRDLSRGNCTYNILDLCLQGIKRLQHFDFDFET
(151) AYRIILFG-ETSWLPFRDLSRGNCTYNILDLCLQGIKRLQHFDFDFET
SC_CDC14 (113) VLQPLAQV-DPPFMPFRDLSRGNCTYNILDLCLQGIKRLQHFDFDFET

      201                                    250
BMY_HPP2_FL (1) -----MGVQPPNRSVILGKAGLALR-----
BMY_HPP2.partial (3) GRRRGQVGAMGVPPNRSVILGKAGLALR-----
HS_CDC14A (150) IDVDEYHYRVRVSGDNCVVEGSHALSGHPKSKI-----ENGWPLHA
HS_CDC14B (200) FNLSEYHYRVRVSGDNCVVEGSHALSGHPKSKI-----ESGKHQS
SC_CDC14 (162) FNLSEYHYRVRVSGDNCVVEGSHALSGHPKSKI-----ESGKHQS

      251                                    300
BMY_HPP2_FL (24) LFAHYQGLDLGLVRHLSSTBERG-PFHSDSCPELTIRRLRFDGCPD
(35) LFAHYQGLDLGLVRHLSSTBERG-PFHSDSCPELTIRRLRFDGCPD
HS_CDC14A (195) PFAHYQGLDLGLVRHLSSTBERG-PFHSDSCPELTIRRLRFDGCPD
(245) PFAHYQGLDLGLVRHLSSTBERG-PFHSDSCPELTIRRLRFDGCPD
SC_CDC14 (211) PFAHYQGLDLGLVRHLSSTBERG-PFHSDSCPELTIRRLRFDGCPD

      301                                    350
BMY_HPP2_FL (73) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
BMY_HPP2.partial (84) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
HS_CDC14A (245) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
HS_CDC14B (295) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
SC_CDC14 (261) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL

      351                                    400
BMY_HPP2_FL (123) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
BMY_HPP2.partial (134) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
HS_CDC14A (292) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
HS_CDC14B (342) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL
SC_CDC14 (311) QLRFRVYHVBENARGLGVCVHCALGSRGTGVLGYLKERCHAGAL

      401                                    450
BMY_HPP2_FL (151) DEKWTQVNPAPLQGXHXSSPPLOAGPDXRGGLYCFVEXKSTFNQGGK
BMY_HPP2.partial (184) DEKWTQVNPAPLQGXHXSSPPLOAGPDXRGGLYCFVEXKSTFNQGGK
HS_CDC14A (335) DEKWTQVNPAPLQGXHXSSPPLOAGPDXRGGLYCFVEXKSTFNQGGK
HS_CDC14B (386) DEKWTQVNPAPLQGXHXSSPPLOAGPDXRGGLYCFVEXKSTFNQGGK
SC_CDC14 (355) DEKWTQVNPAPLQGXHXSSPPLOAGPDXRGGLYCFVEXKSTFNQGGK

```

[illegible][illegible]

Figure 8

```

1 100
BMY_HPP3 (1) ----MARNILPASYDLAYGVVRLTTPVNTYFHFPTCELRQGVV
DM_PRL1 (1) MSITMRQDLAPALILYKGFPTDRSDITLHYIMEAKNNND
HS_PTPCAAX1 (1) ----MARNRPAPEVIVNNMFIDRSTHRTLSTPTEAKKGVV
HS_PTPCAAX2 (1) ----MPEAPEVIVNNMFIDRSTHRTLSTPTEAKKGVV
MM_PTPCAAX (1) ----MNSPAPVLSVNNRPLTTHPTNATLNKPTPELRQGVV
CONSENSUS (1) RNRPAFVLSYKNNRPLTTHNPTNATLNKPTPELRQGVV

51 100
BMY_HPP3 (46) VRVKGIVNIALKRGSTQVDFNFDGTAASSCTDDMEKLMNKEHD
DM_PRL1 (51) VRVCEPTNTDEMTQCTHVKDLANOGTEFPCQDMDSEFFVVLHYQ
HS_PTPCAAX1 (46) VRVGEPTVTTTLVEKEGLHWDAPFDGAPESMDMDSEVRIKREE
HS_PTPCAAX2 (43) VRVCDATYDKFEVEREGHWDAPFDGAPESMDMDNLLNKFREE
MM_PTPCAAX (43) VRVCDATYDKAPVEREGHWDAPFDGAPESMDMDNLLNKFREE
CONSENSUS (51) VRVCDATYD A VEKEGIHLDWPEDDGAPPNQIVDOWL LK KREE

101 150
BMY_HPP3 (96) --PCCVAVHCVVAGCRAPVLVALALIECKKYEDAVQFTRKRRGAFNS
DM_PRL1 (101) QNEEACVAVHCVVAGCRAPVLVALALIECKKYEDAVQFTRKRRGAFNS
HS_PTPCAAX1 (96) --PSCVAVHCVVAGCRAPVLVALALIECKKYEDAVQFTRKRRGAFNS
HS_PTPCAAX2 (93) --PSCVAVHCVVAGCRAPVLVALALIECKKYEDAVQFTRKRRGAFNS
MM_PTPCAAX (93) --PCCVAVHCVVAGCRAPVLVALALIECKKYEDAVQFTRKRRGAFNS
CONSENSUS (101) PGCCVAVHCVVAGLGRAPVLVALALIE GMKYEDAVQFTRKRRGAFNS

151 180
BMY_HPP3 (144) KQLLYLEKYRKMRLRFRDNTGH CCVQ
DM_PRL1 (151) KQLSSLEKYRKMRLRFRDNTGH CCVQ
HS_PTPCAAX1 (144) KQLLYLEKYRKMRLRFRDNTGH CCVQ
HS_PTPCAAX2 (141) KQLLYLEKYRKMRLRFRDNTGH CCVQ
MM_PTPCAAX (141) KQLLYLEKYRKMRLRFRDNTGH CCVQ
CONSENSUS (151) KQLLYLEKYRKMRLRFRDNTGH CCVQ

```

Figure 9A

```

1101                                     1150
BMY_HFP4      (1) -----LRCYEAKS-----
MM_OST-PTP   (1101) RRRKGRPRSEKNGFSUOLMPYNLWRTHRPTHSRRCQSYEAKSRRHOF
RN_PTP-OST    (1101) WRCKGRPRSEKNGFSUOLMPYNLWRTHRPTHSRRCQSYEAKSRAHAQT
CONSENSUS     (1101) R LKGPRSEK GFS ELMYPYNLWRTHRPT HSFRQSYEAKSAAHAQAF

1151                                     1200
BMY_HFP4      (17) ELKPEELKEVKRQOPRLEAEHPANITKNKXPHVLVYDHSRVRLTQLGSE
MM_OST-PTP   (1151) EQEPEELKEVKRQOPRLEAEHPANITKNKXPHVLVYDHSRVRLTQLGSE
RN_PTP-OST    (1151) EQEPEELKEVKRQOPRLEAEHPANITKNKXPHVLVYDHSRVRLTQLGSE
CONSENSUS     (1151) EQEPEELKEVKQDQPRLEAEHPANITKNRYPHVLVYDHSRVRLTQL GSE

1201                                     1250
BMY_HFP4      (66) HSDYINANLVEVYRQEFELACQGLKKITNFWRLVVEQQVHIMLT
MM_OST-PTP   (1201) HSDYINANLVEVYRQEFELACQGLKKITNFWRLVVEQQVHIMLT
RN_PTP-OST    (1201) HSDYINANLVEVYRQEFELACQGLKKITNFWRLVVEQQVHIMLT
CONSENSUS     (1201) HSDYINANLPIGYSHPQEIATQGLKKITLEDVRLVVEQQVHIMLT

1251                                     1300
BMY_HFP4      (115) SMENGRVLCHEHYNLDTSTPTTHGHITTHLLAEPEDEWTREFOQLQV-Q
MM_OST-PTP   (1251) SMENGRVLCHEHYNLDTSTPTTHGHITTHLLAEPEDEWTREFOQLQV-Q
RN_PTP-OST    (1251) SMENGRVLCHEHYNLDTSTPTTHGHITTHLLAEPEDEWTREFOQLQV-Q
CONSENSUS     (1251) SMENGRVLCHEHYNP NSTPTVTHGHITTHLLAEPEDEWTREFOQLQV E

1301                                     1350
BMY_HFP4      (163) QKQRRVQLQFTTIXSDHSVEAPSSLLAFELVQXQVLTQCGVPIVLVHC
MM_OST-PTP   (1301) QKQRRVQLQFTTIXSDHSVEAPSSLLAFELVQXQVLTQCGVPIVLVHC
RN_PTP-OST    (1301) QKQRRVQLQFTTIXSDHSVEAPSSLLAFELVQXQVLTQCGVPIVLVHC
CONSENSUS     (1301) QKQRRVQLQFTTIXSDHSVEAPSSLLAFELVQXQVLTQCGVPIVLVHC

1351                                     1400
BMY_HFP4      (213) RGCCPCGVNRRGTGTFVALSHITQGLEEFVVDVSHAVVLRHMLPMIQT
MM_OST-PTP   (1351) S-----AGRGRTGTFVALSHITQGLEEFVVDVSHAVVLRHMLPMIQT
RN_PTP-OST    (1351) S-----AGRGRTGTFVALSHITQGLEEFVVDVSHAVVLRHMLPMIQT
CONSENSUS     (1351) S AGVGRGTGTFVALLRLRLQLEEEQVVDVFNVTYILRLHRLPMIQT

1401                                     1450
BMY_HFP4      (263) LQYFLHSLCLLNKILEGPHNISSEWPISTWILEQACAKRAANAGFLK
MM_OST-PTP   (1396) LQYFLHSLCLLNKILEGPHNISSEWPISTWILEQACAKRAANAGFLK
RN_PTP-OST    (1396) LQYFLHSLCLLNKILEGPHNISSEWPISTWILEQACAKRAANAGFLK
CONSENSUS     (1401) LSQYFLHSLCLLNKILEGP DASDSGPISTVMDFAQACAKRAANAGFLK

1451                                     1500
BMY_HFP4      (312) EY-----EIKDAEFSAECESEEDPWPYSYDRCCQFSFVSEESSEEDNPL
MM_OST-PTP   (1446) EYKLLQKAIKDETSLLPPDYNNQNSIYS HSQEQFALVVEESP D ML
RN_PTP-OST    (1446) EYKLLQKAIKDETSLLPPDYNNQNSIYS HSQEQFALVVEESP D ML
CONSENSUS     (1451) EYKLLQKAIKDETSLLPPDYNNQNSIYS HSQEQFALVVEESP D ML

1501                                     1550
BMY_HFP4      (357) WKRMICALQGGESGDIPLVTGPGFRELWETVWCHRAVLSVLSVCFENVM
MM_OST-PTP   (1496) ASLFF-----GGFSGRDHVLTGSGFRELWETVWCHRAVLSVLSVCFENVM
RN_PTP-OST    (1496) ASLFF-----GGFSGRDHVLTGSGFRELWETVWCHRAVLSVLSVCFENVM
CONSENSUS     (1501) ASLFF GGFSGRDHVLTGSGFRELWETVWCHRAVLSVLSVCFENVM

1551                                     1600
BMY_HFP4      (407) EK---EKEETMQVETDMVTVAHVSESSEGSWFQDLVETESR-KERE
MM_OST-PTP   (1542) EKQDLVETMQVETDMVTVAHVSESSEGSWFQDLVETESR-KERE
RN_PTP-OST    (1542) EKQDLVETMQVETDMVTVAHVSESSEGSWFQDLVETESR-KERE
CONSENSUS     (1551) EKPDVETMQVETDMVTVAHVSESSEGSWFQDLVETESR-KERE

1601                                     1650
BMY_HFP4      (452) VQLQFPHLPEPELPTLLPFLAVQCCQSGNKKKFGTLLSHSSRGA
MM_OST-PTP   (1591) VQLQFPHLPEPELPTLLPFLAVQCCQSGNKKKFGTLLSHSSRGA
RN_PTP-OST    (1592) VQLQFPHLPEPELPTLLPFLAVQCCQSGNKKKFGTLLSHSSRGA
CONSENSUS     (1601) VQLQFPH ESQ ELFANLTLLTLDVAGQCCSGNKKKFGTLLSHSSR T

```

Figure 9B

```

      1651                                     1700
EMY_HPP4 (502) TQLSTFLAMEQLLQQAGSECTVDLSEVALQSSQACGLMTPTLQYKLYN
MM_OST-PTP (1641) NQLSTFLAMEQLLQQAGSECTVDLSEVALQSSQACGLMTPTLQYKLYN
RN_PTP-OST (1642) NQLSTFLAMEQLLQQAGSECTVDLSEVALQSSQACGLMTPTLQYKLYN
CONSENSUS (1651) NQLGTFLEMEQLLQQAGTERTVDVFNVALKQSQACGLMTPTLEQYIYLYN

      1701                                     1720
EMY_HPP4 (551) CLNSALADLE-----
MM_OST-PTP (1691) CLNSALRNLEERK-----
RN_PTP-OST (1692) CLNSALNLEERAKWPAFC
CONSENSUS (1701) CLNSAL NGLERA K

```

	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

BMV_HPP5	(1)	MAHEMLIGTGVETPERVLLGSGTEKV	100
HS_DSPF8	(1)	MGSDHLEFVYMDAKIASLGGSGGLG	100
MM_NP1	(1)	MGSDHLEFVYMDAKIASLGGSGGLG	100
CONSENSUS	(1)	MAGDRLPRKVMADKIASLLRGPGGLVIDRSFVEYNSS HVLSSVNIC	100
BMV_HPP5	(50)	CSKLVKRRLLQDQVLTITELICGSAKHK	100
HS_DSPF8	(51)	CSKLVKRRLLQDQVLTITELICGSAKHK	100
MM_NP1	(51)	CSKLVKRRLLQDQVLTITELICGSAKHK	100
CONSENSUS	(51)	CSKLVKRRLLQDQVLTITELICGSAKHK	100
BMV_HPP5	(100)	ADSFPLVLTITELICGSAKHK	150
HS_DSPF8	(101)	ADSFPLVLTITELICGSAKHK	150
MM_NP1	(101)	ADSFPLVLTITELICGSAKHK	150
CONSENSUS	(101)	ADSFPLVLTITELICGSAKHK	150
BMV_HPP5	(149)	QPCFLPVPSVGLTRILPHYLGGQDVNLK	200
HS_DSPF8	(151)	QPCFLPVPSVGLTRILPHYLGGQDVNLK	200
MM_NP1	(151)	QPCFLPVPSVGLTRILPHYLGGQDVNLK	200
CONSENSUS	(151)	QPCFLPVPSVGLTRILPHYLGGQDVNLK	200
BMV_HPP5	(199)	DFICESRFRVPINDNYCEKLLPWLKDS	250
HS_DSPF8	(201)	DFICESRFRVPINDNYCEKLLPWLKDS	250
MM_NP1	(201)	DFICESRFRVPINDNYCEKLLPWLKDS	250
CONSENSUS	(201)	DFICESRFRVPINDNYCEKLLPWLKDS	250
BMV_HPP5	(249)	SRSATATAYIMKRMGMSDDAYRFVKRRP	300
HS_DSPF8	(251)	SRSATATAYIMKRMGMSDDAYRFVKRRP	300
MM_NP1	(251)	SRSATATAYIMKRMGMSDDAYRFVKRRP	300
CONSENSUS	(251)	SRSATATAYIMKRMGMSDDAYRFVKRRP	300
BMV_HPP5	(299)	RLKLAALQGD	350
HS_DSPF8	(301)	RLKLAALQGD	350
MM_NP1	(301)	RLKLAALQGD	350
CONSENSUS	(301)	RLKLAALQGD	350
BMV_HPP5	(349)	AAAR	400
HS_DSPF8	(351)	AAAR	400
MM_NP1	(351)	AAAR	400
CONSENSUS	(351)	AAAR	400
BMV_HPP5	(397)	KRSFSLDIKSYVSYSSMAASLHGSSSEDA	450
HS_DSPF8	(391)	KRSFSLDIKSYVSYSSMAASLHGSSSEDA	450
MM_NP1	(393)	KRSFSLDIKSYVSYSSMAASLHGSSSEDA	450
CONSENSUS	(401)	KRSFSLDIKSYVSYSSMAASLHGSSSEDA	450
BMV_HPP5	(447)	PVQELSECTEENSPDKESAPPKLQTL	500
HS_DSPF8	(423)	PVQELSECTEENSPDKESAPPKLQTL	500
MM_NP1	(429)	PVQELSECTEENSPDKESAPPKLQTL	500
CONSENSUS	(451)	PVQELSECTEENSPDKESAPPKLQTL	500
BMV_HPP5	(497)	RSFSLPDRSSGSGVENVYHTFLFLSS	550
HS_DSPF8	(463)	RSFSLPDRSSGSGVENVYHTFLFLSS	550
MM_NP1	(464)	RSFSLPDRSSGSGVENVYHTFLFLSS	550
CONSENSUS	(497)	RSFSLPDRSSGSGVENVYHTFLFLSS	550

Figure 10B

```

551                                     600
BMY_HPP5 (546) PQTSPTSLTSSMIDATESSHFYSASAIYGGSAQFSAMSCSQTETC-----
HS_DSPP8 (513) DSEPTSPDGFPCSEEG-----ACGACGVLIAPFRRAGRECF-----
MM_NPP1 (514) DSEPTSPDGFPCSEEG-----ACGPEAVFSAEYVSAGALEPGNS
CONSENSUS (551) DSPGTPSPDGFPCFSPEG      A  GGA FSAFGR  AP FG

601                                     650
BMY_HPP5 (592) -----DOVYEVVEROKPS-----DRA
HS_DSPP8 (552) -----EGDLRRSEAARA-----EPF
MM_NPP1 (556) SSSGGGGGGGGGGGGGGGGGGSSSSSSSSSSSSSSSSSSSSSSDLRR
CONSENSUS (601) GGS S RR      S      RR

651                                     700
BMY_HPP5 (608) QKRRSMDISPFKQKRRSCOMETCSEIENSRSEELIKVSSSFSFG
HS_DSPP8 (568) QKRRSMDISPFKQKRRSCOMETCSEIENSRSEELIKVSSSFSFG
MM_NPP1 (606) QKRRSMDISPFKQKRRSCOMETCSEIENSRSEELIKVSSSFSFG
CONSENSUS (651) D RTGWPEEPA E QKRRSCOMETEEMVEGRARGELLAALGKQ SFSG

701
BMY_HPP5 (658) SVEIEVS
HS_DSPP8 (618) SVEIEVS
MM_NPP1 (656) SVEIEVS
CONSENSUS (701) SVEIEVS

```

Figure 11

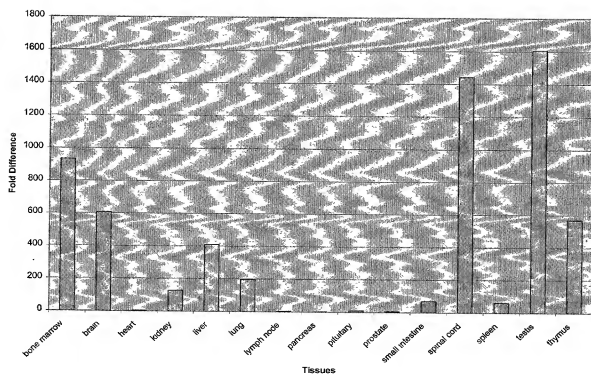


Figure 12.

BMV_HPP5

<u>Protein</u>	<u>Genbank ID</u>	<u>Identities</u>	<u>Similarities</u>
human dual specificity phosphatase 8	gi NP_004411	46%	58%
mouse neuronal tyrosine/threonine phosphatase 1	gi NP_032774	43%	56%

RET31

<u>Protein</u>	<u>Genbank ID</u>	<u>Identities</u>	<u>Similarities</u>
human protein-tyrosine phosphatase DUS8 protein	gi U27193	50.3%	56.8%
the human dual specificity MAP kinase DUSP6 protein	gi AB013382	36.5%	48.3%
human map kinase phosphatase MKP-5 protein	gi AB026436	34.3%	47.2%
mouse RET31 protein	N/A	90%	92%

mRET31

<u>Protein</u>	<u>Genbank ID</u>	<u>Identities</u>	<u>Similarities</u>
human protein-tyrosine phosphatase DUS8 protein	gi U27193	48.5%	55.7%
the human dual specificity MAP kinase DUSP6 protein	gi AB013382	37.4%	49.7%
human map kinase phosphatase MKP-5 protein	gi AB026436	35.2%	46.9%
human RET31 protein	N/A	90%	92%

Figure 13A

1 GAAAAGAAGACGAGGAGGAGCGACGGGACGGGACGCGAGCGGGAGCGCAGCCGCCCTC 60
 61 TCGGCTCCGCGCGCGCGCCCTCGCAAGTCCGGGAGCGAGGGGGCCCCAGGGGAGACGCC 120
 121 GTGACAACTTTCGTTTCCCTCTGAGGGAATTGGGAGGTCGGCGGCCCAAAAGCTTTCAG 180
 181 TCCAGTGTAAAGCTGTTGGAGCGCGGAGCAAGGTAAAGATGATGTAATGCGCTGGCT 240
 241 GCTCCAAAGCATCTTTTGTGTGGAATGGTTATTCAGTCATCTCTTTATGAATCAAATG 300
 301 TGAGGGGCTGCTTTGTGACGGAGTCCTTTGCAAGAGCACATCAACGGGAAAGAGAAAGA 360
 361 GACATTCACCTTGGAGGGCTCTTGCTGAAATGGGTTTAACTCTCTCTTTGCCAGTCACCA 420
 421 CCAGCCTGACCTCATACACTTTTAGTACAATGGAGTGGCTGAGCCTTTGAGCACACCACC 480
 481 ATTACATCATCGTGGCAAATTAAAGAAGAGGTGGGAAAAGAGGACTTATTGTTGTCATG 540
 1 M 1
 541 GCCCATGAGATGATTGGAACCTCAAATTGTTACTGAGAGGTTGGTGGCTCTGCTGGAAGT 600
 2 A H E M I G T Q I V T E R L V A L L E S 21
 601 GGAACGGAAAAAGTGTCTGCTAATTGATAGCCGGCCATTTGTGGAATACAATACATCCAC 660
 22 G T E K V L L I D S R P F V E Y N T S H 41
 661 ATTTTGAAGCCATTAAATCAACTGCTCCAAGCTTATGAAGCGAAGGTTGCAACAGGAC 720
 42 I L E A I N I N C S K L M K R R L Q Q D 61
 721 AAAGTGTTAATTACAGAGCTCATCCAGCATTAGCGAAACATAAGTTGACATTGATTGC 780
 62 K V L I T E L I Q H S A K H K V D I D C 81
 781 AGTCAGAAGTTGTAGTTTACGATCAAAGCTCCCAAGATGTTGCCTCTCTCTCTCAGAC 840
 82 S Q K V V V Y D Q S S Q D V A S L S S D 101
 841 TGTCTTCTACTGTACTCTGGGTAACTGGAGAAGAGCTTCAACTCTGTTCACTCTGCTT 900
 102 C F L T V L L G K L E K S F N S V H L L 121
 901 GCAGGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGGAAAATCCACT 960
 122 A G G F A E F S R C F P G L C E G K S T 141

Figure 13B

961 CTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTTGCCAACATGGGCCAACCCGA 1020
 142 L V P T C I S Q P C L P V A N I G P T R 161

1021 ATTCTTCCCAATCTTTATCTTGGCTGCCAGCGAGATGTCTCAACAAGGAGCTGATACAG 1080
 162 I L P N L Y L G C Q R D V L N K E L I Q 181

1081 CAGAATGGGATGGTTATGTGTTAAATGCCAGCTATACCTGTGCCAAAGCCTGACTTTATC 1140
 182 Q N G I G Y V L N A S Y T C P K P D F I 201

1141 CCCGAGTCTCATTTCCTGCGTGTGCCTGTGAATGACAGCTTTTGTGAGAAAATTTTGCCG 1200
 202 P E S H F L R V P V N D S F C E K I L P 221

1201 TGTTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAGCCTCCAATGGATGTGTTCTA 1260
 222 W L D K S V D F I E K A K A S N G C V L 241

1261 GTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCCTACATCATGAAG 1320
 242 V H L A G I S R S A T I A I A Y I M K 261

1321 AGGATGGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACTATA 1380
 262 R M D M S L D E A Y R F V K E K R P T I 281

1381 TCTCCAACTTCAATTTTCTGGGCCAACTCTCGACTATGAGAAGAAGATTAAGAACCCAG 1440
 282 S P N F N F L G O L L D Y E K K I K N Q 301

1441 ACTGGAGCATCAGGGCCAAAGAGCAAACCTCAAGCTGCTGCACCTGGAGAGCCAAATGAA 1500
 302 T G A S G P K S K L K L L H L E K P N E 321

1501 CCTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAAGCGAGACGCCCTCAGTCCACCCTGT 1560
 322 P V P A V S E G G Q K S E T P L S P P C 341

1561 GCCGACTCTGCTACCTCAGAGGCAGCAGGACAAGGCCCGTGCATCCCGCCAGCGTGCC 1620
 342 A D S A T S E A A G Q R P V H P A S V P 361

1621 AGCGTGCCACGCGTCAGCCGTCGCTGTAGAGGACAGCCCGTGGTACAGGCGCTCAGT 1680
 362 S V P S V Q P S L L E D S P L V Q A L S 381

1681 GGGCTGCACCTGTCCGACAGAGGCTGGAAGACAGCAATAAGCTCAAGCGTTCTCTCT 1740
 382 G L H L S A D R L E D S N K L K R S F S 401

1741 CTGGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTACATGGCTTCTCC 1800
 402 L D I K S V S Y S A S M A A S L H G F S 421

1801 TCATCAGAAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGATGGGACCAACAAG 1860
 422 S S E D A L E Y Y K P S T T L D G T N K 441

Figure 13C

1861 CTATGCCAGTTCTCCCTGTTTCAGGAACATATCGGAGCAGACTCCCGAAACCAGTCCTGAT 1920
 442 L C Q F S P V Q E L S E Q T P E T S P D 461

1921 AAGGAGGAAGCCAGCATCCCAAGAAGCTGCAGACCGCCAGGCCTTCAGACAGCCAGAGC 1980
 462 K E E A S I P K K L Q T A R P S D S Q S 481

1981 AAGCGATTGCATTTCGGTCAGAACCCAGCAGCAGTGGCACCGCCAGAGGTCCCTTTTATCT 2040
 482 K R L H S V R T S S S G T A Q R S L L S 501

2041 CCACTGCATCGAAGTGGGAGCGTGGAGGACAATTACCACACCAGCTTCCTTTTCGGCCTT 2100
 502 P L H R S G S V E D N Y H T S F L F G L 521

2101 TCCACCGGCCAGCAGCACCTCACGAAGTCTGCTGGCCTGGGCCTTAAGGCTGGCACTCG 2160
 522 S T S Q Q H L T K S A G L G L K G W H S 541

2161 GATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAGCAGCTGGTATTTGCCACA 2220
 542 D I L A P Q T S T P S L T S S W Y F A T 561

2221 GAGTCTCACACTTCTACTCTGCCTCAGCCATCTACGGAGGCAGTGCAGTTACTCTGCC 2280
 562 E S S H F Y S A S A I Y G G S A S Y S A 581

2281 TACAGCTGCAGCCAGCTGCCACTTGCAGAGACCAAGTCTATTCTGTGCGCAGGCGGCAG 2340
 582 Y S C S Q L P T C G D Q V Y S V R R R Q 601

2341 AAGCCAAGTGACAGAGCTGACTCGCGGCGGAGCTGGCATGAAGAGAGCCCCCTTTGAAAAG 2400
 602 K P S D R A D S R R S W H E E S P F E K 621

2401 CAGTTTAAACGCAGAAGCTGCCAAATGGAATTGGAGAGAGCATCATGTTCAGAGAACAGG 2460
 622 Q F K R R S C Q M E F G E S I M S E N R 641

2461 TCACGGGAAGAGCTGGGGAAAGTGGGCAGTCAGTCTAGCTTTTCGGGCAGCATGGAATC 2520
 642 S R E E L G K V G S Q S S F S G S M E I 661

2521 ATTGAGGCTCCTGAGAAGAAAGACACTTGTGACTTCTATAGACAATTTTTTTTCTTG 2580
 662 I E V S 665

2581 TTCACAAAAAATCCCTGTAAATCTGAAATATATATATGTACATACATATATATTTTG 2640

2641 GAAAAATGGAGCTATGGTGTAAAAGCAACAGTGGATCAACCCAGTTGTTACTCTCTTAAC 2700

2701 ATCTGCATTTGAGAGATCAGCTAATACTTCTCTCAACAAAAATGGAAGGCAGATGCTAG 2760

2761 AATCCCCCTAGACGGAGGAAAACCATTTTATTTCAGTGAATTACACATCCTCTTGTCTT 2820

Figure 13D

2821 AAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCCTACCATTTTACGTTGTGCT 2880
 2881 ACTAAGAGATCTCAAATATTAGTCTTTGTCCGACCCTTCCATAGTACACCTTAGCGCTG 2940
 2941 AGACTGAGCCAGCTTGGGGTCAGGTAGGTAGACCTGTTAGGGACAGAGCCTAGTGGTA 3000
 3001 AATCCAAGAGAAATGATCCTATCCAAAGCTGATTACAAAACCCACGCTCACCTGACAGCC 3060
 3061 GAGGGACACGAGCATCACTCTGCTGGACGGACCATTAGGGGCCTTGCCAAGGTCTACCTT 3120
 3121 AGAGCAAACCCAGTACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCATATCTGGTA 3180
 3181 GCCCATTTTCTAGGCATTGTGAATAGGTAGGTAGCTAGTCACACITTTTACAGCAATTC 3240
 3241 AACTGTCATGACAAAATCCCGTGGGCCTAGATGGAGATAATTTTTTTTCTCTCAG 3300
 3301 CTTTATGAGAGAGGAACTGTCTAGGATTGAGCTGACACCAGCAACCTGGCAACA 3360
 3361 TCACGATTTAAGCTAAGGTTGGGAGGCTAACGAGTCTACCTCCCTCTTTGTAAATCAAAG 3420
 3421 AATTGTTTAAATGGGATTGTCAATCCTTTAAATAAAGATGAAGTGGTTTCAAGCCAAA 3480
 3481 TGTGAATTTATTTGGGTTGGTAGCAGAGCAGCAGCACCTTCAAATTTCTAGCCAAAGCAG 3540
 3541 ATGTTTTTGCCCTTTCTGCTTCACTGCATGGATACAGTTGGTAAATGTAATAATATGGC 3600
 3601 AGAATTTTATAGGAACTTCTAGGGAGGTAAATTTATGGGAAGATTAAAGAAAGGTACAAA 3660
 3661 TTGCTGAGGAGAAGCAGGAAACCTGTTTCCTTAGTGGCTTTTATCCCCTCGGCATGCGAT 3720
 3721 GGGGCTGATGTTTCTATGATTGCCTCAGACTTTCACATTTACTAGTAGGGCTGAGAGAGG 3780
 3781 CTTTAGTGAGGAAGGAATATTCAGAATAAAACGGTTGAGAAAGCTGAGAAGACCATTGAG 3840
 3841 TTTTGATCAGTTGTGAATAGAGTGCAAGCCATGGCCAAGCTGTTTTTGGAAACGCTGGC 3900
 3901 CGGCGTGTCTTCACTGAAAAAGCAAATCAAATGGAGCGAGAGCAAAGGGGCGTCTCTCA 3960

Figure 13E

3961 GTCCTCAACCTACAATCACTGTATGGAATCGGTCCTGGCAGCTGAACATAGGAGGTCAC[•] 4020
 4021 GGAACAAGTGATAGTGCAGATTGGCTTTCAAACATCCTCCTGGCTTGAGTTTATCAGCT[•] 4080
 4081 ACAATGTGGGTCTCTTTTGAAGCCTTAATTCACAACAGCAGCTTTTGGGGGTGGGGCT[•] 4140
 4141 GGGGGGTGTGTGTCATTGTTCTTTCCCTTCCTGTAAGTGTGCTAGTTGCTGCCTCGTAT[•] 4200
 4201 CTCAGGTTTTTCTCTGTTTTTGAGAAATGGACAGTTTTTTGACCAGGATGTGACTTCATG[•] 4260
 4261 TTTCTATGGTGACTTCTAAAACCAGCACAGAATGATATGACTCAACACAGACCGACTTG[•] 4320
 4321 GTTATGGGGATGATGAGCCGCACAGACCTCACTAGTTGTGCACAAATAATGTGCTATGAT[•] 4380
 4381 GGGGTGTAAAGTGAAGCAGAAGAGGGTCAGCCGCATTGTTATGATACTGGGAAAGTGCT[•] 4440
 4441 GGTCACGATTTGAGTTAGTTTTTAGATATACATTGAAATCTTTAATCAGACATTCTCAA[•] 4500
 4501 GTTTCACACAGTAGTTTTTGATGTTATGTACACACACACCAAAATGTGAACAGTTCACCA[•] 4560
 4561 CTTCACAGAGTGTGGTCATGCCCAAAACATGTTTAAGAAAGGAAAGCAGTAGCTCCTTGCT[•] 4620
 4621 AACGATGTTTCAGGAGGTTTGGGGCACTTGGTTTTTAATGAGCTTCTGTCAATTTAGGGCTT[•] 4680
 4681 CTCTTGGCCATGGTCCCTTCTCTCTGGAACGTGATGTAGTCACATCCTACAGCCTTTA[•] 4740
 4741 GTGCTGGTTCAGTAGTGTGAGATAATCAGTTCTTGGAATCGAGACTGCCCTGGCGAAGGG[•] 4800
 4801 GTGGCCTCGGAGGCAGGCTCTGGAGCTGCTTGGATGTCCTTTAGGTGGGGTGGTGGCTGGC[•] 4860
 4861 TCTCTTCAGCATGTAATTGGGGAAACCTTCGCGTCTACTAGGGGTGATACAGATGGTGAT[•] 4920
 4921 TTAAAGAGCAAAACAGACTTCTATGTGAGAAGTGTGGAAATGATTTAGGACATGTG[•] 4980
 4981 TAAAGTTAGATGAAAGACTGTAATGTTTAATATGAATATAGTGTTCTTTTGAAGTAA[•] 5040
 5041 GCCAGCTGTGAAACGGTTAAACTGTGCATTCTCATTTTGATGTGTCATGTATGTTAATG[•] 5100
 5101 TATGAAATGATTAAATAAAATCAAACCTGGTACCTGTTTATACATAAATACGAGAAAAGA[•] 5160

Figure 13F

5161 CCTATCTTTGCAGCCATAAACTCGGTGGGAACACCACCACTCAAGTTGCCAAAGGAGGCA 5220
5221 GTGGTGAAACCTGTCCTGTTCTCACTTAAATGAGGATTTAGCTCAAAATAAAGTGGTGGT 5280
5281 GTCATCAGGTTTATTCCGTGTTCTGTCATTCACATGGAACACCGGATGATTAGCTAACAG 5340
5341 TTTAGTGCCAGCCTTCATTCTTTACTGTGTACGTTAAATGCACACTACAGTGAAAAAGCC 5400
5401 TAAGACACTTGGTAAATAATTTCTAGCTGACTGATTCAGAACACACAAG 5450

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Figure 14A

		1	50
RET31	(1)	-----	-----
mRET31	(1)	-----	-----
DUS6	(1)	-----	-----
DUS8	(1)	-----	-----
MKP-5	(1)	MPPSPLDDDR VVVALSRPVRQDNLNCLDSSYLGSANPGSNSHPVVIATTV	
		51	100
RET31	(1)	-----	-----
mRET31	(1)	-----	-----
DUS6	(1)	-----	-----
DUS8	(1)	-----	-----
MKP-5	(51)	VSLKAANLTYMPSSSGSARSNLNCGCSSASCTVATYDKDNQAQTQAIAG	
		101	150
RET31	(1)	-----MAHEMSTGVTERLVALLSECT-----	
mRET31	(1)	-----MAHEMSTGVTERLVALLSECT-----	
DUS6	(1)	-----MIDTLRPVPFASDAISKTAWEHQELQCN-----	
DUS8	(1)	-----MAGDRPRKVMDAKKAALGLRGG-----	
MKP-5	(101)	TTTTAIGTSTTCFANQVANNIENIGSPSSSGVSSPVSGTPKQLASIKII	
		151	200
RET31	(24)	-----GRLLIDSRPFVEYNHSHILPAININCSKLM-----	
mRET31	(24)	-----GRLLIDSRPFVEYNHSHILPAININCSKLM-----	
DUS6	(32)	-----GRLLIDSRQEQLYESHIESAINWALGIM-----	
DUS8	(25)	-----GGPLVIDSSEFVEYNHSHISSNNTCSKLM-----	
MKP-5	(151)	YPNDLAKKMTKCSKSHLPSGPWIDCRPFVEYNHSHIQGVAINCNRKI	
		201	250
RET31	(55)	KRRLOQD--KVITELTHQSTKHKVDIDCNQVVVYDSS--QVVASLSSD	
mRET31	(55)	KRRLOQD--KVITELTHQSTKHKVDIDCNQVVVYDSS--QVVGSLSSD	
DUS6	(63)	PRRLOAGNLPVRALFTRGETRRRTTRCGTDTVVLYDSSSDWNENTGEH	
DUS8	(56)	KRRLOQG--KVITELTQPARSOVATEPCDVVVYDQSI--RQASVLQAD	
MKP-5	(201)	SRRLQGGKITVLDLISCREKDSFRIRIFSKETVYDENTNEPSRVVPSQ	
		251	300
RET31	(102)	CELVYLLGKLRKNSNSVHLIAGGFAPESRCFPGLCCEGH--STLIVETCTSOE	
mRET31	(102)	CELVYLLGKLRKNSNSVHLIAGGFAPESRCFPGLCCEGH--STLIVETCTSOE	
DUS6	(113)	SHGLLKKLRDEGCRAFYEEGFKKQALFFSLHPTET--LDGSCSSSE	
DUS8	(103)	SEALILSKTQGCDFSVATITGGFTSSCFPGLCCEGPAAALFMS--SOE	
MKP-5	(251)	--FLHIVYTESLKEEGKEPLMKGGSSSRQNTIELCDNHLQLQCREMVGGG	
		301	350
RET31	(151)	CLPVANIC-----	PTRI
mRET31	(151)	CLPVANIC-----	PTRI
DUS6	(161)	FLPVLLGLLRISDSSSDIESDLDRDPNSATDSGSPLSNSQPSFPLFI	
DUS8	(153)	CLPVDSVC-----	LPRI
MKP-5	(300)	ASAASSLLPQPIPTTFDIEN-----	AEIPLI
		351	400
RET31	(163)	LENLYLGCQDVLNLEIQQNGIGYVLNASYTCF--KPDFTFESSFETRVVF	
mRET31	(163)	LENLYLGCQDVLNKLMOQNGIGYVLNASNTCF--KPDFTFESSFETRVVF	
DUS6	(211)	LEPPLYLGCQDSTHLEWVDEEGTSTYLENNIPNLENLFNAGEFKMKKTFIT	
DUS8	(165)	LEPPLYLGCQDVLNKLDMQNGIYVLNASNSCF--KPDFTFCESREVRVPE	
MKP-5	(326)	LEPPLYLGCNECDQDLNLMQNLNIGYVNNVTHLEFLYHYEKGLFNMKRIEPL	

Figure 14B

```

401                                     450
RET31 (212) NDSKCEKLLPWLDKSNFTIEKAKNSNGCVLVHCLAGISRSATIAIAYIMF
mRET31 (212) NDSKCEKLLPWLDKSNFTIEKAKNSNGCVLVHCLAGISRSATIAIAYIMK
DUS6 (261) SDHNSQNLSCFFPEALSFDEARGKNGCVLVHCLAGISRSATIAIAYIMK
DUS8 (214) NDSKCEKLLPWLDKSNFTIEKAKNSNGCVLVHCLAGISRSATIAIAYIMK
MKP-5 (376) TDSNKNQNTROYEENFTIEEAKHQCKGLLECOQVRSRSATIAIAYIMK

451                                     500
RET31 (262) RQDMSLEDAYRFVKKRFTISPNNFNGLDLDYKTKIKNNQMSGPKSKL
mRET31 (262) RQDMSLEDAYRFVKKRFTISPNNFNMGQLDYKTKIKNNQMSGPKSKL
DUS6 (311) KLNLSNNDAYDVRKKKEISPNNFNMGQLDERTILGLSSPCDNRVPAQ
DUS8 (264) RQDMSDDAYRFVKKRFTISPNNFNGLDLDYKTKIKLLAALQGDGP--
MKP-5 (426) RTRNNNDAYRFVKKRFTISPNNFNMGQLDERTILKNGNIPRLTPEKL

501                                     550
RET31 (312) KTHLEKPNEDVPAVSESGQKSETPISPFCDNSTSEAAKQRP--VHPAS
mRET31 (312) KTHLEKPNEDVPAVSESGQKKSALTSPECANSTSEASGQRL--VHPAS
DUS6 (361) QLYFTTPSNQNYQVSLQST-----
DUS8 (312) -TPSG----TCEPSPNAGAPLPDPPTSESAAGNNPREGGLSANG
MKP-5 (476) MGVEYTVV-----

551                                     600
RET31 (360) VESVPSVQESLLEDSPVQVATSGELSLSKRLDPSNHLKRSFSLDIKSVSY
mRET31 (359) VE---RLQESLLEDSPVQVATSGELSLSKRLDPSNHLKRSFSLDIKSVSY
DUS6 (382) -----
DUS8 (357) EL---PAPNPPATSAALQCFRELHLSNRLQNNHLKRSFSLDIKSA-Y
MKP-5 (483) -----

601                                     650
RET31 (410) SASMAASLHGFSSSEDELEYYSFSTTLDGNTNKLQFSPVQELSEQSEETS
mRET31 (406) SASMAASLHGFSS-EDLDYKPSATLDGNTNKLQFSPVQVELSEQSEETS
DUS6 (382) -----
DUS8 (403) APSRRPDGFPDPDGFAPKLCQLDSPSG-----AALGESPSPDS
MKP-5 (483) -----

651                                     700
RET31 (460) PDKEEASIPKKLQTAHSDSQSKRLHSVRTSSSGTAQRSLLSLHRS-
mRET31 (455) PDKEEAHIEQPPPRISESQVTRLHSVRTGSSGSTQRPFFSLHRS-
DUS6 (382) -----
DUS8 (443) PDAAPEARPPRRRRPFP-----AGSPARSPASLSLN
MKP-5 (483) -----

701                                     750
RET31 (509) VENNYHTSFLFSLSTSQQHLTKSA--LGLKWHSDTLFOTSTPSLTSS
mRET31 (504) VENNYHTNLFGLSTSQQHLTKSA--LGLKWHSDTLFQSSAPSLTSS
DUS6 (382) -----
DUS8 (476) FGDAAARQTFRHSLSALSAFGLPGFQPAQFGAMAPPDSE--GTSPDGP
MKP-5 (483) -----

751                                     800
RET31 (557) MYEATSSHFYSASIIYCSASYSAHSCSQLETCGDQVYSVRRQKPSR
mRET31 (552) MYEATPSHLYSASIIYCNSSYSAGCGQLTCSGDQIYSVRRQKPTSR
DUS6 (382) -----
DUS8 (524) WQNSPQ-----GQQAGVLTAPAGRAGAGPGGGGDLRRREAAARAP
MKP-5 (483) -----

```

Figure 14C

RET31 (607) A P S R S M H P S P F P C F K R R S C Q M E F C S T I M S E N R S R S D E L I G K W C S S S F S
mRET31 (602) A P S R S M H P S P F P C F K R R S C Q M E F C S T I M S E N R S R S D E L I G K W C S S S F S
DUS6 (382) -----
DUS8 (567) R P A R T G M P P E P A P F P C F K R R S C Q M E F C S T I M Y E G R A R G S D E L I A L C H S S F S
MKP-5 (483) -----

851

RET31 (657) G S M F I I E V S
mRET31 (652) G S M F I I E V S
DUS6 (382) -----
DUS8 (617) G S M F I I E V S
MKP-5 (483) -----

Figure 15.

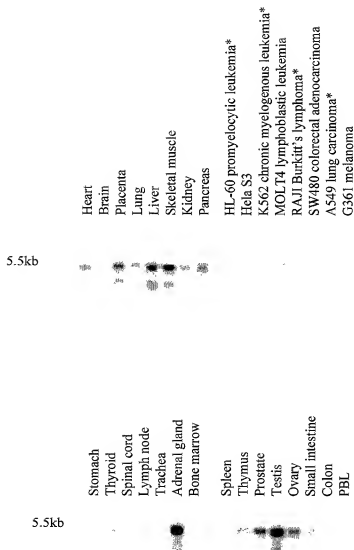


Figure 16A

1 GACTGAGGTTGTCAGCCCGAGTGTAAGCTGTTGGAGTGAGGGCAGAAAGGTAAAGGATGA 60
 61 TGTAATGCCTGGCTGCCCTAGAGCATCTTTTGTGTGGGATGGGTATTTCCCATCATCTCT 120
 121 ATGAATCTAGTGTGAGGGGCTGCTTTGTGGAAGGAATCCTTTGCAAGAGCATATCAACAG 180
 181 GAAAGAGAAAAGAGACATTTCAGTTGGAGGGCTCTTGTGTAATGGATTAACTCTCTCTCT 240
 241 GCCAGTCACCACTAGCCTGACCTCATACATTTTTAGTACAATGGAGTGGCTGAGCCTTTG 300
 301 AGCACAGCACCATTACATCATCGTGGCAAATTAAAGAACGAGGTGGGGAAAGAGGACTTA 360
 361 TTGTTGTCATGGCCCATGAGATGATTGGAAGTCAAATTTGTTACTGAGAGCTTGGTGGCTC 420
 1 M A H E M I G T Q I V T E S L V A L 18
 421 TGCTGGAAGTGAACGGAAAAAGTGTCTAATTGATAGCCGACCATTTGTGGAATACA 480
 19 L E S G T E K V L L I D S R P F V E Y N 38
 481 ATACGCTCTCACATTTTGAAGCCATTAAATATCAACTGCTCCAAACTGATGAAGCGAAGGT 540
 39 T S H I L E A I N I N C S K L M K R R L 58
 541 TGCAACAGGACAAAGTATTAAATTACAGAACTAATCCACCAATCTACAAAGCATAAAGTTG 600
 59 Q Q D K V L I T E L I H Q S T K H K V D 78
 601 ACATTGACTGCAATCAAAGAGTGGTAGTTTATGATCACAGTTCACAAAGATGTTGGTTCTC 660
 79 I D C N Q R V V V Y D H S S Q D V G S L 98
 661 TGTCGTCAGACTGCTTTCTCACTGTACTTCTGGGTAAGCTGGAGAGAAGCTTCAACTCTG 720
 99 S S D C F L T V L L G K L E R S F N S V 118
 721 TCCACCTGCTTGCAGGTGGCTTTGCTGAGTTCTCTCGTTGTTCCCTGGCCTCTGTGAAG 780
 119 H L L A G G F A E F S R C F P G L C E G 138
 781 GAAAGTCACTCTAGTCCCTACCTGCATATCTCAGCCTTGCTTACCTGTTGCGAACATTG 840
 139 K S T L V P T C I S Q P C L P V A N I G 158
 841 GGCCAACTCGAATTCTTCCCAATCTCTATCTTGGCTGCCAGCGAGATGTCCTCAACAAGG 900
 159 E T R I L P N L Y L G C Q R D V L N K D 178
 901 ACCTGATGCAACAGAATGGGATTGGCTATGTGTTAAATGCCAGCAATACCTGTCCAAAGC 960
 179 L M Q Q N G I G Y V L N A S N T C P K P 198

Figure 16B

961	CTGACTTCATACCTGAATCTCACTTCCTGCGAGTGCCTGTGAATGACAGCTTTTGTGAGA	1020
199	<u>D F I P E S H F L R V P V N D S F C E K</u>	218
1021	AAATCCTACCATGGTTGGACAAGTCTGTGGATTTCATTGAGAAAGCAAAAGCCTCCAATG	1080
219	<u>I L P W L D K S V D F I E K A K A S N G</u>	238
1081	GCTGTGTGCTTATCCACTGCTTAGCTGGGATCTCTCGCTCCGCCACTATTGCTATTGCTCT	1140
239	<u>C V L I H C L A G I S R S A T I A I A Y</u>	258
1141	ACATCATGAAGAGGATGGACATGCTCTTAGATGAGGCTTACAGATTGTGAAAGAAAAA	1200
259	<u>I M K R M D M S L D E A Y R F V K E K R</u>	278
1201	GACCTACTATATCTCCGAATTTTAATTTATGGGCCAACTCATGGACTATGAGAAGACGA	1260
279	<u>P T I S P N F N F M G Q L M D Y E K T I</u>	298
1261	TTAATAACCAGACTGGAATGTCAGGGCCAAAGAGCAAACTGAAGCTGCTGCACCTAGACA	1320
299	<u>N N Q T G M S G P K S K L K L L H L D K</u>	318
1321	AACCCAGTGAGCCCGTGCCTGCAGCCCTCAGAGGGCGGATGGAAGAGTGCACCTGCTCTCA	1380
319	<u>P S E P V P A A S E G G W K S A L S L S</u>	338
1381	GTCCACCCTGTGCCAACTCGACCTCGGAGGCATCAGGGCAAGGCTTGTGCATCCTGCAA	1440
339	<u>P P C A N S T S E A S G Q R L V H P A S</u>	358
1441	GTGTGCCCGCTTACAGCCGCTCACTCTTAGAGGACAGTCCGCTGGTACAGGCGCTCAGTG	1500
359	<u>V P R L Q P S L L E D S P L V Q A L S G</u>	378
1501	GGCTCCAGCTGTCTCAGAGAAGCTGGAAGACAGCACTAAGCTCAAGCGTTCTTCTCTC	1560
379	<u>L Q L S S E K L E D S T K L K R S F S L</u>	398
1561	TCGATATCAAATCTGTTTCAATATTCAGCCAGTATGGCCGCGTCCCTACAGGGCTTCTCGT	1620
399	<u>D I K S V S Y S A S M A A S L H G F S S</u>	418
1621	CAGAGGAGGCTTTAGACTACTGCAACCTTCTGCCACACTGGATGGGACCAACAAGCTCT	1680
419	<u>E E A L D Y C K P S A T L D G T N K L C</u>	438
1681	GCCAGTTCTCCCCGTTTCAGGAGGTATCAGAACAGAGTCCAGAGACCAGCCCGGATAGG	1740
439	<u>Q F S P V Q E V S E Q S P E T S P D K E</u>	458
1741	AGGAGGCCACATCCCCAAGCAGCCCCAACCTCCAGGCCTTCTGAGAGCCAGGTACAC	1800
459	<u>E A H I P K Q P Q P P R P S E S Q V T R</u>	478
1801	GCTTGCACTCAGTGAGAACCAGGAGTAGTGCGGTCCACCCAGAGGCCCTTCTCTCGCCAC	1860
479	<u>L H S V R T G S S G S T Q R P F F S P L</u>	498

Figure 16C

1861	TGCATCGGAGCGGGAGTGTAGAGGACAATTACCATACCAACTTTCCTTTTGGGCTTTCCA	1920
499	H R S G S V E D N Y H T N F L F G L S T	518
1921	CCAGCCAGCAACACCTCACCAAGTCTGCAGGGCTTGGCCTCAAGGGCTGGCACTCAGATA	1980
519	S Q Q H L T K S A G L G L K G W H S D I	538
1981	TTCTGGCTCCCACTCCTCTGCCCCCTCCCTGACCAGCAGTTGGTATTTTGCCTACGGAGC	2040
539	L A P Q S S A P S L T S S W Y F A T E P	558
2041	CTTCACACTTGTACTCTGCTTCAGCCATCTATGGAGGCAACAGCAGTTACTCTGCCTACA	2100
559	S H L Y S A S A I Y G G N S S Y S A Y S	578
2101	GCTGTGGCCAGCTGCCCACTTGCAGTGACCAAACTCTATTCTGTTCGTAGGCGGCAGAAAGC	2160
579	C G Q L P T C S D Q I Y S V R R R Q K P	598
2161	CTACTGACAGAGCTGACTCGAGGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAGCAGT	2220
599	T D R A D S R R S W H E E S P F E K Q F	618
2221	TTAAACGCAGAAGCTGCCAAATGGAATTTGGAGAGAGCATTATGTCGGAGAACAGGTCCA	2280
619	K R R S C Q M E F G E S I M S E N R S R	638
2281	GGGAGGAGCTGGGCAAGTGGGCGAGCCAGTCCAGCTTCTCCGGCAGCATGGAGATCATCG	2340
639	E E L G K V G S Q S S F S G S M E I I E	658
2341	AGGTCTCTTGAGAAGACCTCGTCGCTTCTGTTGACAGTTTTGTTTCTCTGTTCACAAAAAA	2400
659	V S	660
2401	TAGTCCCTGTAAATCTGAAATATGTATATGTACATACATATATATTTTTGGAAATATAGAG	2460
2461	CTACGGTATAAAAGCAACAGATGGATCAACACAGTTGTTCTCTCAGCACCTGCCTGAGA	2520
2521	ATAGCTAACTCTCAGAAAAGATTGGAAGGGTAGATGTTAGAAATCTCCAGCCAGGAGAA	2580
2581	GAGATTTGGTTCAGTGAATTGCACATCTTCTTGTTCCTACAAAAGCAAGGGTTTGTITG	2640
2641	TITGTATGTGTTTGTGTTTTAATGTTAGAGGGCAAAATCCCTCCCATTTTCACGTGCAAC	2700
2701	AGAGGTCTCAGAACTCATCTCTGTCCAGGCCTTCCCTAGTGCACCTTAGCGCTAA	2756

Figure 17.

Ret31_DSpc G P T R N V C Q R V L N K E L M Q N G S G V L A S Y T C K P D F I P S H F
 DUS8_DSpc G L I R H V S Q K D V L N K D L M T Q N S S V L A S N S C K P D F I C S R F
 MKP5_DSpc E L P F F L N E Q T A Q D L D T M O R L N G V I V T T H L L Y H Y E K G L F N Y
 DUS6_DSpc F P V E F C A K E S T N L D V L E E F S K I L V T P N L N L F E N A G F K Y

 Ret31_DSpc L V V N S F C E K I L P W L D K S V D E K K A S N G C V L V L A S S A T A
 DUS8_DSpc M V I N N Y C E K I L P W L D K S I E D K K L S S C Q V L V L A S S A T A
 DUS6_DSpc K I I S H W S Q N S Q F F E A I S D E R G K N C G V L V L A S S A T A
 MKP5_DSpc K L A T S N K Q N R O Y F E E A F E E H Q C G K L I I Q V V A T V

 Ret31_DSpc L I K R M D S L D E T R E K R T S P F M L R D L D Y K K
 DUS8_DSpc A I K T M G M S D D R E D R S S P F M L R D L D Y K K
 MKP5_DSpc A L A H T R M T M T S A K F G K R I S P F M L R D L D Y K K
 DUS6_DSpc A L L Q K L N L S M N D A D T C M K K S N S P F M L R D L D Y K K

Figure 18

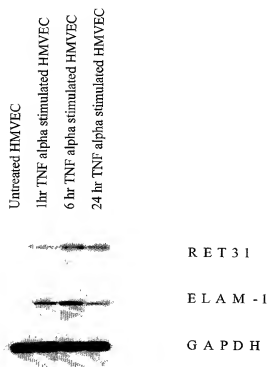


Figure 19A

1 GAAAGAAGACGAGGAGGAGCGACGGGACGGGAGCGGAGCGCAGCCGCCCTC 60
 61 TCGGCTCCGCGCGCGGCCCTCGCAAGTCCGGGAGGCGAGGGGGCCCCAGGGGAGACGCC 120
 121 GTGACAACTTTCGTTCCCTCTGAGGGAATTGGGAGGTCGGCGGCCCAAAAGCTTTCAG 180
 181 TCCAGTGTAAGCTGTTGGAGCGCGGGAGCAAAGGTAAAGAATGATGTAATGCGCTGGCT 240
 241 GCTCCAAAGCATCTTTTGTGTGGAATGGTTATTCCAGTCATCTCTTTATGAATCAAATG 300
 301 TGAGGGGCTGCTTTGTGGACGGAGTCCTTTGCAAGAGCACATCAACGGGAAAGAGAAAAGA 360
 361 GACATTCACTTGGAGGGCTCTTGCTGAAAATGGGTTTAACTCTCCTTTTGCCAGTCACCA 420
 421 CCAGCCTGACCTCATACACTTTTAGTACAATGGAGTGGCTTGAGCCTTGAGCACACCACC 480
 481 ATTACATCATCGTGGCAAATTAAAGAAGGAGGTGGGAAAAGAGGACTTATTGTTGTCATG 540
 1 M 1
 541 GCCCATGAGATGATTGGAACCTCAAATTGTTACTGAGAGGTGGTGGCTCTGCTGGAAAGT 600
 2 A H E M I G T Q I V T E R L V A L L E S 21
 601 GGAACGGA AAAAGTGTCTGCTAATTGATAGCCGGCCATTTGTGGAATACAATACATCCCAC 660
 22 G T E K V L L I D S R P F V E Y N T S H 41
 661 ATTTTGAAGCCATTATATCAACTGCTCCAAGCTTATGAAGCGAAGGTTGCAACAGGAC 720
 42 I L E A I N I N C S K L M K R R L Q Q D 61
 721 AAAGTGTTAATTACAGAGCTCATCCAGCATTACGCGAAACATAAGGTTGACATTGATTGC 780
 62 K V L I T E L I Q H S A K H K V D I D C 81
 781 AGTCAGAAGTTGTAGTTTACGATCAAAGCTCCCAAGATGTGCTCTCTCTCTTCAGAC 840
 82 S Q K V V V Y D Q S S Q D V A S L S S D 101
 841 TGTTTCTCACTGTACTTCTGGGTAAACTGGAGAAGAGCTTCAACTCTGTTCACTCGCTT 900
 102 C F L T V L L G K L E K S F N S V H L L 121
 901 GCAGTGGGTTTGCTGAGTTCTCTCGTTGTTCCCTGGCCTCTGTGAAGGAAAATCCACT 960
 122 A G G F A E F S R C F P G L C E G K S T 141

Figure 19B

961	CTAGTCCCTACCTGCATTTCAGCCTTGCTTACCTGTTGCCAACATTGGGCCAACCCGA	1020
142	L V P T C I S Q P C L P V A N I G P T R	161
1021	ATTCTTCCCAATCTTTATCTTGGCTGCCAGCGAGATGTCCTCAACAAGGAGCTGATACAG	1080
162	I L P N L Y L G C Q R D V L N K E L I Q	181
1081	CAGAATGGGATGGTTATGTGTTAAATGCCAGCTATACCTGTCCAAAGCCTGACTTTATC	1140
182	Q N G I G Y V L N A S Y T C P K P D F I	201
1141	CCCGAGTCTCATTTCCTGCGTGTGCCTGTGAATGACAGCTTTTGTGAGAAAATTTTGCCG	1200
202	P E S H F L R V P V N D S F C E K I L P	221
1201	TGGTTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAGCCTCCAATGGATGTGTCTA	1260
222	W L D K S V D F I E K A K A S N G C V L	241
1261	GTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCCTACATCATGAAG	1320
242	V H C L A G I S R S A T I A I A Y I M K	261
1321	AGGATGGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAAGACCTACTATA	1380
262	R M D M S L D E A Y R F V K E K R P T I	281
1381	TCCCAAACTTCAATTTTCTGGGCCAACTCCTGGACTATGAGAAGAAGATTAGAACCAG	1440
282	S P N F N F L G Q L L D Y E K K I K N Q	301
1441	ACITGGAGCATCAGGGCCAAAGAGCAAACCTCAAGCTGCTGCACCTGGAGAAGCCAAATGAA	1500
302	T G A S G P K S K L K L L H L E K P N E	321
1501	CCTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAAGCGAGACGCCCTCAGTCCACCCTGT	1560
322	P V P A V S E G G Q K S E T P L S P P C	341
1561	GCCGACTCTGCTACCTCAGAGGCAGCAGGACAAAGGCCCGTGTCATCCCGCCAGCGTGCC	1620
342	A D S A T S E A A G Q R P V H P A S V P	361
1621	AGCGTGCCCGCGTGCAGCCGTCGCTGTTAGAGGACAGCCCGCTGGTACAGGCGCTCAGT	1680
362	S V P S V Q P S L L E D S P L V Q A L S	381
1681	GGGTGCACCTGTCCGACAGAGGCTGGAAGACAGCAATAAGCTCAAGCGTTCTCTCTCT	1740
382	G L H L S A D R L E D S N K L K R S F S	401
1741	CTGGATATCAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTACATGGCTTCTCC	1800
402	L D I K S V S Y S A S M A A S L H G F S	421
1801	TCATCAGAAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGATGGGACCAACAAG	1860
422	S S E D A L E Y Y K P S T T L D G T N K	441

100215335442004

Figure 19C

1861	CTATGCCAGTTCCTCCCCTGTTTCAGGAACCTATCGGAGCAGACTCCCGAAACCAGTCCTGAT	1920
442	L C Q F S P V Q E L S E Q T P E T S P D	461
1921	AAGGAGGAAGCCAGCATCCCCAAGAAGCTGCAGACCGCCAGGCCTTCAGACAGCCAGAGC	1980
462	K E E A S I P K K L Q T A R P S D S Q S	481
1981	AAGCGATTGCATTTCGGTCAGAACCCAGCAGCAGTGGCACCGCCAGAGGTCCTTTTATCT	2040
482	K R L H S V R T S S S G T A Q R S L L S	501
2041	CCACTGCATCGAAGTGGGAGCGTGGAGGACAATTACCACACCAGCTTCCTTTTCGGCCTT	2100
502	P L H R S G S V E D N Y H T S F L F G L	521
2101	TCCACCAGCCAGCAGCACCTCAGCAAGTCTGCTGGCCTGGGCCTTAAGGGCTGGCACTCG	2160
522	S T S Q Q H L T K S A G L G L K G W H S	541
2161	GATATCTTGGCCCCCAGACCTCTACCCCTTCCTGACCAGCAGCTGGTATTTTGCCACA	2220
542	D I L A P Q T S T P S L T S S W Y F A T	561
2221	GAGTCTCAGCACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCAGTTACTCTGCC	2280
562	E S S H F Y S A S A I Y G G S A S Y S A	581
2281	TACAGCTGCAGCCAGCTGCCACTTGCAGAGACCAAGTCTATTCTGTGCCAGGCGGCAG	2340
582	Y S C S Q L P T C G D Q V Y S V R R R Q	601
2341	AAGCCAAGTGACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAG	2400
602	K P S D R A D S R R S W H E E S P F E K	621
2401	CAGTTTAAACGCAGAAGCTGCCAAATGGAATTGGAGAGAGCATCATGTACAGAAACAGG	2460
622	Q F K R R S C Q M E F G E S I M S E N R	641
2461	TCACGGGAAGAGCTGGGGAAAGTGGGCAGTCAGTCTAGCTTTTCGGGCAGCATGGAATC	2520
642	S R E E L G K V G S Q S S F S G S M E I	661
2521	ATTGAGGTCCTCTGAGAAGAAAGACACTTGTGACTTCTATAGACAATTTTTTTTTCTTG	2580
662	I E V S	665
2581	TTCACAAAAAAATTCCTGTAAATCTGAAATATATATATGTACATACATATATATTTTTTG	2640
2641	GAAATGGAGCTATGGTGTAAAAGCAACAGGTGGATCAACCCAGTTGTTACTCTCTTAAC	2700
2701	ATCTGCATTTCAGAGATCAGCTAATACTTCTCTCAACAAAAATGGAAGGGCAGATGCTAG	2760
2761	AATCCCCCTAGACGGAGGAAAACCATTTTATTTCAGTGAATTACACATCCTCTTGTTCTT	2820

Figure 19D

2821 AAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTACCATTTTACGTTGTGCT 2880
 2881 ACTAAGAGATCTCAAATATTAGTCTTTGTCCGGACCCTTCCATAGTACACCTTAGCGCTG 2940
 2941 AGACTGAGCCAGCTTGGGGGTGAGGTAGGTAGACCTGTTAGGGACAGAGCCTAGTGGTA 3000
 3001 AATCCAAGAGAAATGATCCTATCCAAAGCTGATTACAAAACCCACGCTACCTGACAGCC 3060
 3061 GAGGGACACGAGCATCACTCTGCTGGACGGACCATTAGGGGCTTGCCAAGGTCTACCTT 3120
 3121 AGAGCAAACCCAGTACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCATATCTGGTA 3180
 3181 GCCCATTTTCTAGGCATTGTGAATAGGTAGGTAGCTAGTCACACTTTTCAGACCAATTCA 3240
 3241 AACTGTCTATGCACAAAATTCCTGGGGCTAGATGGAGATAATTTTTTTCTCTCAG 3300
 3301 CTTTATGAAGAGAAGGGAACTGTCTAGGATTCAGCTGAACACCAGGAACCTGGCAACA 3360
 3361 TCACGATTTAAGCTAAGGTTGGGAGGCTAACGAGTCTACCTCCCTCTTTGTAATCAAG 3420
 3421 AATTGTTTAAATGGGATTGTCAATCCTTTAAATAAAGATGAACCTGGTTTCAAGCCAAA 3480
 3481 TGTGAATTTATTTGGGTTGGTAGCAGAGCAGCAGCACCTTCAAATTCTCAGCCAAAGCAG 3540
 3541 ATGTTTTTGCCTTTCTGCTTCACTGCATGGATACAGTTGGTAAATGTAATAATATGGC 3600
 3601 AGAATTTTATAGGAAACTTCCTAGGGAGGTAAATATGGGAAGATTAAAGAAAGGTACAAA 3660
 3661 TTGCTGAGGAGAAGCAGGAAACCTGTTTCCTTAGTGGCTTTTATCCCTCGGCATGCGAT 3720
 3721 GGGGCTGATGTTCTATGATTGCCTCAGACTTTCACATTTACTAGTAGGGCTGAGAGAGG 3780
 3781 CTTTAGTGAGGAAGGAATATTGAGAATAAAACGGTTGAGAAAGCTGAGAAGACCATTGAG 3840
 3841 TTTTGATCAGTTGTGAATAGAGTGCAAAGCCATGGCCAAGCTGTTTTTGGAAACGCTGGC 3900
 3901 CGGCGTGTCTTCAGTGGAAGAAAGCAAATCAAATGGAGCGAGAGCAAAGGGCGCTCCTCA 3960
 3961 GTCTCAACCTACAATCACTGTATGGAATCGGTCTGGCAGCTGAACATAGGAGGTCACT 4020

Figure 19E

4021 GGAACAAGT⁺GATAGTGCAGATTGGCTTTCAAACATCCTCTGGCTTGAGTTTATCAGCT⁺ 4080
 4081 ACAATGTGGGTCCTCTTTTGAAGCCTTAATTCACAACAGCAGCTTTTGGGGGTGGGGCT⁺ 4140
 4141 GGGCGGGTGTGTGTCATTGTCTTTCCCTTCTGCTAAGTGTGCTAGTTGCTGCCTCGTAT⁺ 4200
 4201 CTCAGGTTTTTCTCTGTTTTTTGAGAAATGGACAGTTTTTTGACCAGGATGTGACTTCATG⁺ 4260
 4261 TTTCCCTATGGTGACTTCTAAAACCAGCACAGAATGATATGACTCAACACAGACCGACTTG⁺ 4320
 4321 GTTATGGGATGATGAGCCGCACAGACCTCACTAGTTGTGCACAAATAATGTGCTATGAT⁺ 4380
 4381 GGGGTGTAAAGTGAAGCAGAAGAGGGTCAGCCGCATTGTTATGATACTGGGAAATGCT⁺ 4440
 4441 GGTCAACGATTTGAGTTAGTTTTTTAGATATACATTGAAATCTTTAATCAGACATTCTCAA⁺ 4500
 4501 GTTTCACACAGTAGTTTTTGATGTTATGTACACACACACAAATGTGTAACAGTTCACCA⁺ 4560
 4561 CTTCAGAGTGTGGTCATGCCAAAACATGTTTAAGAAAGGAAAGCAGTAGCTCCTTGCT⁺ 4620
 4621 AACGATGTTTCAGGAGGTTTGGGGCACTTGGTTTTAATGAGCTTCTGTCAATTAGGGCTT⁺ 4680
 4681 CTCTTGGCCATGGTCCCCTTCTTCTGGAACGTGATGTAGTCACATCCTACAGCCTTTA⁺ 4740
 4741 GTGCTGGTTCAGTGTGTCAGATAATCAGTTCTTGGAATCGAGACTGCCGTGGCGAAGGG⁺ 4800
 4801 GTGGCTCGGAGGCAGGCTCTGGAGCTGCTTGGATGTCTTTAGGTGGGGTGGTGGCTGGC⁺ 4860
 4861 TCTCTTCAGCATGTAATTGGGGAAACCCTCGCGTCTACTAGGGGTGATACAGATGGTGAT⁺ 4920
 4921 TTTAAAGAGCAAAACTAGACTTCTATGTGAGAAGTGTCTGGAAAATGATTTAGGACATGTG⁺ 4980
 4981 TAAAGTTAGATGAAAGACTGTAAATGTTTAATATGAATATAGTGTCTTTTGAAGTAAG⁺ 5040
 5041 GCCAGCTGTTGAACGTTAAACTGTGCATTTCTCATTTTGATGTGTCATGTATGTTAATG⁺ 5100
 5101 TATGAAATGATTAAATAAAATCAAACCTGGTACCTGTTTATACATAAATACGAGAAAAGA⁺ 5160

Figure 19F

5161 CCTATCTTTGCAGCCATAAACTCGGTGGGAACACCACCACTCAAGTTGCCAAAGGAGGCA 5220
5221 GTGGTGAAACCTGTCTCTCACTTAAATGAGGATTTAGCTCAAAATAAAGTGGTGGT 5280
5281 GTCATCAGGTTTATTCCGTGTTCTGTCATTACATGGAACACCGGATGATTAGCTAACAG 5340
5341 TTTAGTGCCAGCCTTCATTCTTTACTGTGTACGTTAAATGCACACTACAGTAAAAAGCC 5400
5401 TAAGACACTTGGTAAATATTTTCTAGCTGACTGATTCCAGAACACACAAG 5450

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Figures 20A

1 CCACCGCTCCGGCTCTTGCTCCAGTGCCATGCAGGTGCAGGATGCAACAGSCGGCCC 60
 61 TCAGCGGTGGCTTCCTCAGCTCCTTTCTCCAGGCCGCCGCACTCCACCTCAGACCCA 120
 121 GTACTGCGGTGCAGCAGGCCCGCGGGGCTCTGGCTGGGCTCCGGCTCTGCCACGAAG 180
 181 CTGTGTCTCTGCTCCTCTCTCCAGTGATGGTGGCTGTTTCTCAGTCAGCCATGCAGAG 240
 241 GGAACCCAACTTTCCCCGAAAGAAAAAGAAATTTAGAACGTCCAACACCAAGTACACA 300
 301 AAGTAGGGGAGCGTTTACGSCATGTCATTCTGGACACATGGCATGTTCCATGGCGTGT 360
 361 GCGGTAGAGCTTGCAAGTATGAGAACCCAGCCGCTGGAGTGAGCAGGAGCAAGCCATT 420
 421 AAGGGGGTTTACTCATCTGGGTCAGTGATAATATACTGCCCATGGCCCCCATCTCTCT 480
 481 GAGCTCCTGGAGAAGTACCACATCATTTGATCAGTTCTCAGCCATGGCATAAAACAATA 540
 541 ATCAACCTCCAGCGCCCTGGTGAGCATGCTAGCTGTGGGAACCTCTGGAACAAGAAGT 600
 601 GGCTTCACATACCTTCTGAGGCTTTTCATGGAGGCTGGCATTCTACTTCTACAAATTCGGA 660
 1 M E A G I Y F Y N F G 11
 661 TGAAGGATTATGGGTAGCGTCTCTTACTACTATCCTAGATATGGTGAAGGTGATGACA 720
 12 W K D Y G V A S L T T I L D M V K V M T 31
 721 TTTGCCTTACAGGAAGGAAAAAGTAGCTATCCATTGTGTCAGGGCTTGGTCGAACAGGT 780
 32 F A L Q E G K V A I H C H A G G R G G 51
 781 GTTTTAATAGCTGTTACTTAGTTTTCGAACGAGAATGACTGCTGACCAAGCAATTATA 840
 52 V L I A C Y L V F A T R M T A D Q A I I 71
 841 TTTGTGCGGSCAAAGCGACCCCAATTCATACAAACAGAGGACAGCTCCTCTGTGTAAGG 900
 72 F V R A K R P N S I Q T R G Q L L C V R 91
 901 GAATTTACTCAGTTTCTAACTCCTCTCCGCAATATATCTCTCTGCTGTGATCCCAAAGCA 960
 92 E F T Q F L T P L R N I F S C C D P K A 111
 961 CATGTGTCACTTACCTCAATATCTAATTCGCCAGCGTCATCTGCTTCATGGTTATGAG 1020
 112 H A V T L P Q Y L I R Q R H L L H G Y E 131
 1021 GCACGACTTCTGAAACAGTGCACAAAATATCCACCTAGTTTGCAAATTCGCTGCTGAC 1080
 132 A R L L K H V P K I I H L V C K L L L D 151
 1081 TTAGCGGAGAACAGGCCAGTGATGATGAAGGATGTGTCGAAGGACCTGGTCTCTCTGCT 1140
 152 L A E N R P V M M K D V S E G P G L S A 171

Figures 20B

1141 GAAATAGAAAAGACAATGTTCTGAGATGGTCACCATGGAGTAAAGAGTTACTGAGG 1200
172 E I E K T M S E M V T M Q L D K E L L R 191

1201 CATGACAGTGATGTGTCCAACCGCCTAACCCCATGCGAGTGGCAGCAGATTTTGACAA 1260
192 H D S D V S N P P N P T A V A A D F D N 211

1261 CGAGGCATGATTTTCTCAATGAGCAACAGTTTGACCTCTTTGGAAAAGCGGAATGTT 1320
212 R G M I F S N E Q Q F D P L W K R R N V 231

1321 GAGTGCCTTCAACCCCTGACTCATCTGAAAAGCGGCTCAGCTACAGTGACACAGATTTA 1380
232 E C L Q P L T H L K R R L S Y S D S D L 251

1381 AAGAGGCGGAGAACCTCCTGGAGCAAGGGGAGACTCCACAGACAGTGCTGCCAGATC 1440
252 K R A E N L L E Q G E T P Q T V P A Q I 271

1441 TTGGTTGGCCACAGCCAGCCAGCAGAAGCTCATAAGCCATTGTTACATCCACAGTCT 1500
272 L V G H K P R Q Q K L I S H C Y I P Q S 291

1501 CCAGAACAGACTTTACACAAGGAAGCCTTGGTTGCGAGCACACTTTCTTCTGAGGTGAG 1560
292 P E P D L H K E A L V R S T L S F W S Q 311

1561 TCAAAGTTTGGAGGCCTGGAAGGACTCAAAGATAATGGGTACCAATTTCCATGGAAGG 1620
312 S K F G G L E G L K D N G S P I F H G R 331

1621 ATCATTCCAAAGGAAGCACAGCAGAGTGGAGCTTTCTCTGCAGATGTTTCAGGCTCACAC 1680
332 I I P K E A Q Q S G A F S A D V S G S H 351

1681 AGCCCTGGGGAGCCAGTTTCAACCAGCTTTGCAAAATGTCATAAGGATCCAAACCTGCT 1740
352 S P G E P V S P S F A N V H K D P N P A 371

1741 CACCAGCAAGTGCTCTCACTGTGAGTGTAAAACATGGTGTGGGAGCCCTGGCTCTGTC 1800
372 H Q Q V S H C Q C K T H G V G S P G S V 391

1801 AGGCAGAACAGCAGGACACCCGGAAGCCCTCTGGACTGTGGCTCCAGTCCCAAGCACAG 1860
392 R Q N S R T P R S P L D C G S S P K A Q 411

1861 TTCTTGGTTGAACATGAAACCCAGGACAGTAAAGATCTGTCTGAAGCAGCTTCACTCT 1920
412 F L V E H E T Q D S K D L S E A A S H S 431

1921 GCATTACAGTCTGAATTGAGTGTGAGGCAAGAAGATACTGGCGGCCAAGCCCTAGCA 1980
432 A L Q S E L S A E A R R I L A A K A L A 451

1981 AATTTTAATGAATCTGTAGAAAAGGAGGAACATAAAGGAAGGTAGAAATGTGGCAGAAA 2040
452 N L N E S V E K E E L K R K V E M W Q K 471

2041 GAGCTTAATTCCCGAGATGGAGCTTGGGAAAGAATATGTGGGAGAGGGAACCTTTCATC 2100
472 E L N S R D G A W E R I C G E R D P F I 491

Figures 20C

2101 CTATGCAGCCTTGATGTGGTCTTGGGTGGAGCACTGAAGGAGCCTGTAATCACCAGAAG 2160
 492 L C S L M W S W V E Q L K E P V I T K E 511
 2161 GATGTGGACATGTTGGTTGACAGCGAGCAGATGCCCGCAGAAGCACTTTTTTATTAGAG 2220
 512 D V D M L V D R R A D A A E A L F L L E 531
 2221 AAGGGACAGCACCAGACTATTCTCTGCGTGTGCACTGCATAGTGAACCTGCAGACAATT 2280
 532 K G Q H Q T I L C V L H C I V N L Q T I 551
 2281 CCGTGGATGTGGAGGAAGCTTTCCTTGCCCATGCCATTAAAGCATTCTACTAAGGTTAAT 2340
 552 P V D V E E A F L A H A I K A F T K V N 571
 2341 TTTGATCTGAAAATGGACCAACAGTTTACAACACCCGAAAGAAAATATTTAAGCACAG 2400
 572 F D S E N G P T V Y N T L K K I F K H T 591
 2401 CTGGGAAGAAAAAGAAAAATGACAAAAGATGGCCCTAAGCCCTGGCCTTAGCTTTCACCTC 2460
 592 L E E K R K M T K D G P K P G L * 607
 2461 ATGGTGAAATATTTTCAGACCTAAAGATCCAGATAGTATCTCTGTTCATATGTGAATAAGTT 2520
 2521 GAAGATTGTGGGGCTACTTTTTCTCATAGCACTTTATTTTGAATGTTGTAGTTTGTGCT 2580
 2581 GAGAATGTCGTCOGTATTGGAACCAATTATTTATTTTAAATATATTTAAGCTACATTT 2640
 2641 TTGTTTTGAAAAATTGCCATAAATTTGGTGCCACTTCTCTTTATTTATTGACTGAGTTA 2700
 2701 ATATTATTGTATTAACATTTTAAAGTATATGGTGTTCACATTTCTTATTTCTTTTGACATTT 2760
 2761 TGGAAATAATCATAACTTGTCTTTCCAAAATAACCAATTTCTTGATGGAACCTCTTCCTAG 2820
 2821 AGTTTTTACCAAATAGCTAACTTTAGTAGTAAACCTCATTTGTGTATCATTCGCCOCACA 2880
 2881 GATGAACTAAGAAAGTCACCAAGTGTCTTAAGCTGTTTATATTTGTTACGAAGAAGGCT 2940
 2941 ATTGCTACAATATTTTAAAGGTTCTTTTTTAACCTTGAAATTTTTGTGTTTTCTCTTT 3000
 3001 CTTTTTATAAATGTAAACAGAGGGTTTCAAAGCATATTATTTTTCAGAGAGATTAGTTTT 3060
 3061 ACTTTAATGAGTGACTGTGAAGTGGTTGGGATTTTTTGTCTGTAGAAAGTAGACTTGCT 3120
 3121 CTTTGTGAGATTTCCAAACAACCTTGCCAGCCTTGCTGTCAAAGGAGGAGGAGCAGT 3180
 3181 TCTCAACACACCAAGCCTTATTCCTACTCCCTTGGGTGCTGCTGAGCCAAATAGCATCT 3240
 3241 TTACAGAGGAAGTGGGATCAGAGGCAGGAAGTGTGGAAGTTGCTAAGAAGCAGGGCTTG 3300

Figures 20D

3301 CCTCTGCTCTCCCGGGGACTCCACAGGGATATTCTGTGACGGGCAGGGGCTCTGTGCCAGC 3360
 3361 COTGCTCTCTCAGATGCCACAGCCACTCTGCAGAGGTGACTCTTGGAGCTGGAGGAAGTC 3420
 3421 AAAACTGGGCCACTGTTTGTACTGATGGTGTATTAGCATGAGCAGCGTGGCCCTGGCCCC 3480
 3481 ACACGCCAAATCTGCCACTCCATAGACCCACTTGCCTCAAGGCTTTATATTGGCTGTCT 3540
 3541 TTCTTACAATGAGAATTAAGATTTTAAACTGAAGTTGACCATACAGGTGTCATTAGCCC 3600
 3601 TAACTGGCTTCATGTAAGAAGGGTGACTGCCTAAACTAGTTCCTTGTAACTGAACCATC 3660
 3661 AATTATCAGTTGAAGCCATACATTTTATTTAAATTAATATACGTAGATACCAGAGGCCAAG 3720
 3721 CCACAGAGAGGATAAATAGTCTCTCCCAATAAAGTGATATTAATCAGACTAATTTCTGAAC 3780
 3781 TAAAGAAGTTACTGCTTAAAGACGGAATTCAGGGGAAGCAAGACTCATTAGAACAAAT 3840
 3841 GAAATTTCTCCAGTCCTACATTTCTGAATTGACTTCTAGCACATCAAAATATTTCAGTC 3900
 3901 ATTATCAGTCTCATTAACGAAATGCCAATGCTAAATGCAGTGTCTTTCACACTGTTT 3960
 3961 TAATTTTCTTGGGAATTGAGTCCAGTGGATGTTAATGGAGTGGGTTGCCCATCCCTGAA 4020
 4021 ATGTCTTATTTTCAAGTGCCTGGCTGGGAAGAAGGGGAAGAACAATTCATTATATC 4080
 4081 CAAAGATACACTATAAAAAATAGAGTTTTTACCAAAAAAAGATGTTTGTCTCATCTCAGT 4140
 4141 AGGCCTCATTTGGGCAAGTGAOCCACAGTCTTTTGGCGAGTTTGCTATTTCCTGCTGTTGA 4200
 4201 AATACTTGTTTCAACTTAGAGAACAGTTATGATGTGACCATAGCATGGCACAACATAAAA 4260
 4261 TCTAAGCCTGAAACCTGAAAAAAGAGATATGACAAGGGAATTAATCAGGCTATACATAA 4320
 4321 GTAATGTATTTTATTTGAATAAAAAATAAAGAGCAACCCATAAAAAAATAAAAAA 4380
 4381 AAAAAAAAAAAG 4393

Figures 21

1 CCACGCGTCCGGCGAGGGGACGCGTGGCGGAGCGGGCTGGCCAGCCTCGGCCCCCATG 60
 61 ACCCGCTGTCTGTGCCCTTTCCACGCGATGGGCGTGAGCCCCCAACTTCTCTGGGT 120
 1 M G V Q P P N F S W V 11
 121 GCTTCCGGGCGGGCTGGCGGACTGGCGCTGCGCGGCTCCCGCCCACTACCACTTCT 180
 11 L P G R L A G L A L P R L P A H Y Q F L 31
 181 GTTGACCTGGGCGTGGGCACTGGTGTCCCTGACGGAGCGGGCCCCCTCACAGCA 240
 31 L D L G V R H L V S L T E R G P P H S D 51
 241 CAGCTGCCCCGGCCTCACCTGCACCGCTGGCGATCCCCGACTTCTGCCCGCGGGCCC 300
 51 S C P G L T L H R L R I P D F C P P A P 71
 301 CGACGAGATCGACCGCTTCGTGCAGATCGTGGACGAGGCCAACGCACGGGAGAGGCTG 360
 71 D Q I D R F V Q I V D E A N A R G E A V 91
 361 GGGAGTGCACCTGTGCTCTGGGCTTTGGCCGCACTGGCACCATGCTGGCCTGTTACCTG 420
 91 G V H C A L G F G R T G T M L A C Y L V 111
 421 GAAGGAGCGGGGCTTGGCTGCAGGAGATGCCATTGCTGAAATCCGACGACTACGACCCGG 480
 111 K E R G L A A G D A I A E I R R L R P G 131
 481 CTCATCGAGACCTATGAGCAGGAGAAAGCAGTCTTCCAGTTCTACGACGAAACGAATA 540
 131 S I E T Y E Q E K A V F Q F Y Q R T K * 150
 541 AGGGGCTTAGTACCTTCTACCAGGCCCTCACTCCCTTCCCCATGTTGTGATGGGGC 600
 601 CAGAGATGAAGGGAAGTGGACTAAAGTATTAAACCTCTAGCTCCCATTTGGCTGAAGACA 660
 661 CTGAAGTAGCCACCCCTGCAGGCAGGTCTGATTGAAGGGGAGGCTTGTACTGCTTTGT 720
 721 TGAATAAATGAGTTTTACGAACCGAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 780
 781 AAA 840
 841 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGC 878

Figure 22

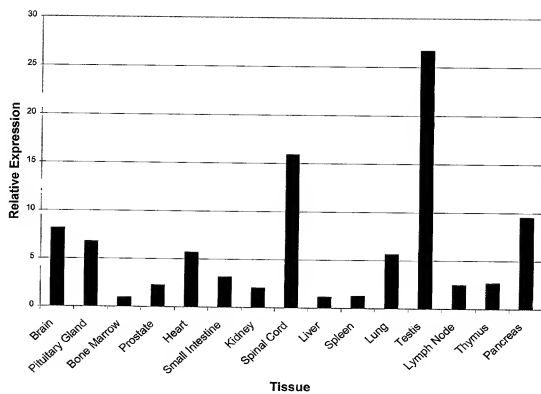


Figure 23

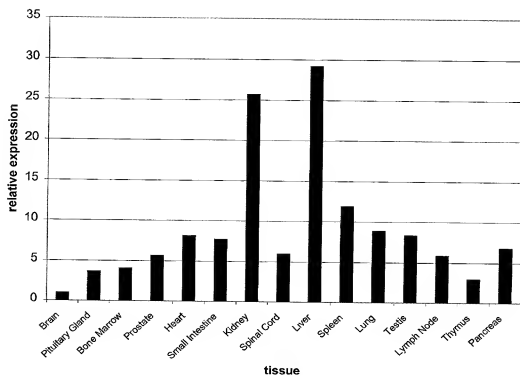


Figure 24.

BMV HPP1

<u>Protein</u>	<u>Genbank ID</u>	<u>Identities</u>	<u>Similarities</u>
human protein tyrosine phosphatase	gi P32587	27%	39.6%
mouse protein tyrosine phosphatase	gi NP_035346	27.9%	40.5%
Schizosacchomyces Pombe protein tyrosine phosphatase PYP3 protein	gi NP_002839	27.5%	36.7%

BMV HPP2

<u>Protein</u>	<u>Genbank ID</u>	<u>Identities</u>	<u>Similarities</u>
human S. cerevisiae CDC14 homolog A	gi NP_003663	33.1%	44.1%
human S. cerevisiae CDC14 homolog B	gi NP_003662	33.1%	45.8%
yeast soluble tyrosine-specific protein phosphatase Cdc14p protein	gi NP_002839	33.1%	45.8%

Figure 25.

RET31			
<u>Protein</u>	<u>Genbank ID</u>	<u>Identities</u>	<u>Similarities</u>
human protein-tyrosine phosphatase DUS8 protein	gi U27193	50.3%	56.8%
the human dual specificity MAP kinase DUSP6 protein	gi AB013382	36.5%	48.3%
human map kinase phosphatase MKP-5 protein	gi AB026436	34.3%	47.2%
mouse RET31 protein	N/A	90%	92%

Relative Abundance of BMV HPP1 in Normal Tissues

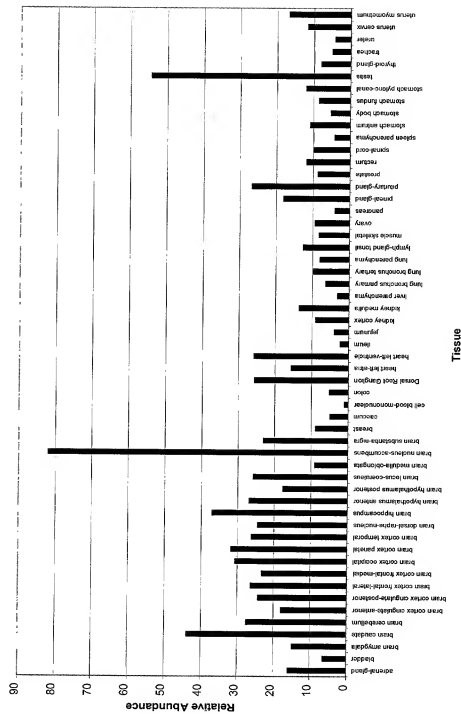
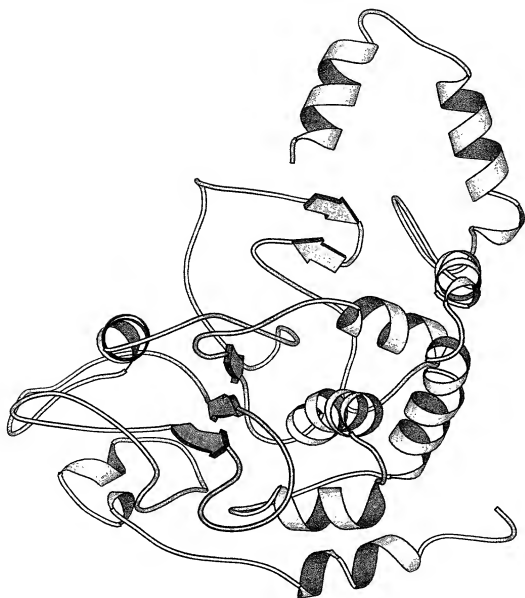


Figure 27

HPP1	MAAGVLPQNE	QPYSTLVNNS	EC.VANMKGN	LERPTPKYTK	39
pdblaax	MEMEKEFEQI	DKSGSWAAY	QDIRHEASDF	PCRVAKLPKN	KNRNRYRDVS	50
HPP1	VGERLRHVIP	GHMCSMACG	GRACKYENPA	RWSEQEQAIK	GVYSSWVTDN	89
pdblaax	PFDH..SRIK	LHQEDNDYIN	ASLIKME...EAQRS	YILTQGPLPN	90
HPP1	ILAMARPSSE	LLEKYHIIDQ	FLSHGIKTII	NLQRPGE..H	ASCGNPLEQE	137
pdblaaxTCGHFWEM	VWEQSRGVV	MLNRVMEKGS	LKCAQYWPQK	128
HPP1	S.....GFTYL	PEAFMEAG..IYFYNEF	158
pdblaax	EEKEMIFEDT	NLKLTLISED	IKSYTIVRQL	ELENLTQET	REILHFHYTT	178
	**			** *	*	
HPP1	WKDYGVA.SL	TTILDMVKVMTFALQE	GKVAIHCHAG	LGRTGVLIAC	203
pdblaax	WPDFGVPESP	ASFLNFLFKV	RESGSLSPFH	GPVVVHSSAG	IGRSGTFCLA	228
HPP1	YLVFATR...MTADQ	AIIFVRAKRP	NSI....QTR	GQLLCVREFT	241
pdblaax	DTCLLLMDKR	KDPSSVDIKK	VLEMRKFRM	GLIQTADQLR	FSYLAVIEGA	278
HPP1	QFLTPLRNIF	SCCDPKAHAV	TLPQYLIRQR	HLLHGYEARL	LKHVPKIIHL	291
pdblaax	KFIM.....	GDSSVQDOWK	ELSHEDLEPP	PGHIPPPRP	312
HPP1	VCKLLLDLAE	NRPVMMKDVS	EGPGLSAEIE	KTMSEMVTMQ	LDKELLRHDS	301
pdblaax	PKRILEPHN.	321

Figure 28



HPP1 Homology Model

Figure 29

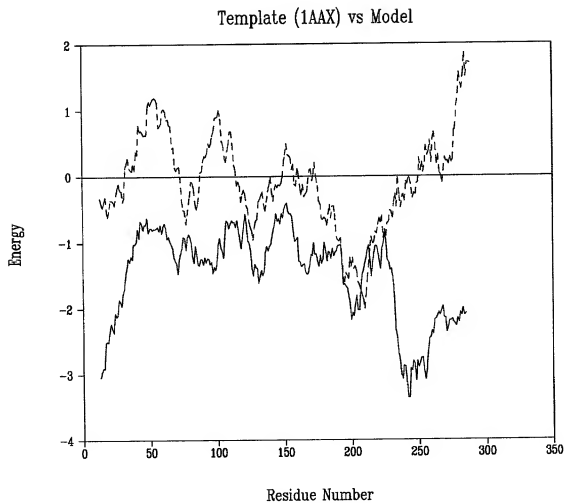


Figure 30

Relative Abundance of BMV_HPP2 in Normal Tissues

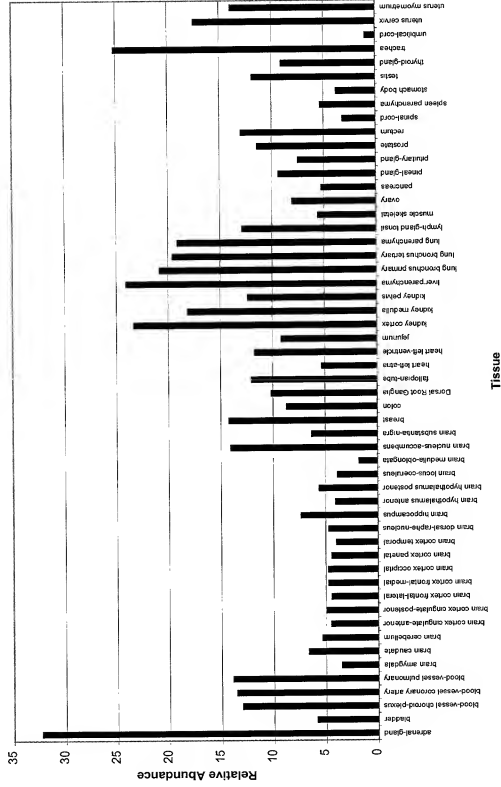


Figure 31

```

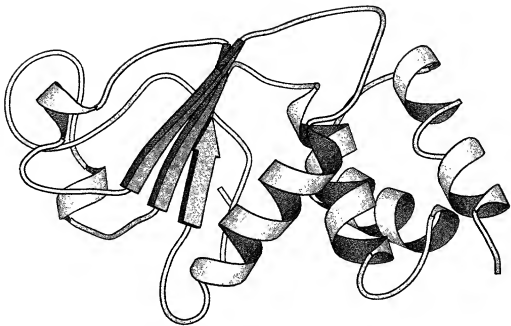
      20      30      40      50      60      70
pdblvrA GCYSLPSQPCNEVTPRIYVGNASVAQDIPKLQKLGITHVLNAAEGRSFMHVNTNANFYKD
      :  ::  ....  ::  ::  ::
BMY_HPP2 MGVQPPNFSWVLPGRLAGLALPRLPAHYQFLDLGVRLVSLTE-RGPPHSDSCP-----
      10      20      30      40      50

      80      90      100      110      120      130
pdblvrA SGITYLGIKANDTQEFNLSA--YFERAADFTDQALAQKNGRVLVHCREGYSRSPTLVIAY
      ::  ..  :  :  :  ::  .....  :  :  :  ::  :  :  :  :  :  :
BMY_HPP2 -GLTLHRLRIPD---FCPPAPDQIDRFVQIVDEANARGEA-VGVHICALGFGRGTGTLACY
      60      70      80      90      100

      140      150      160      170      180
pdblvrA LMMRQKMDVKSALSIVRQNREIGPNDFLAQLCQLNDRLAKEGKLP
      :  .  .  .  :  :  :  :
BMY_HPP2 LVKERGLAAGDAIAEIRRLRPGSIETYEQEKAVFQFYQRTK
      110      120      130      140      150

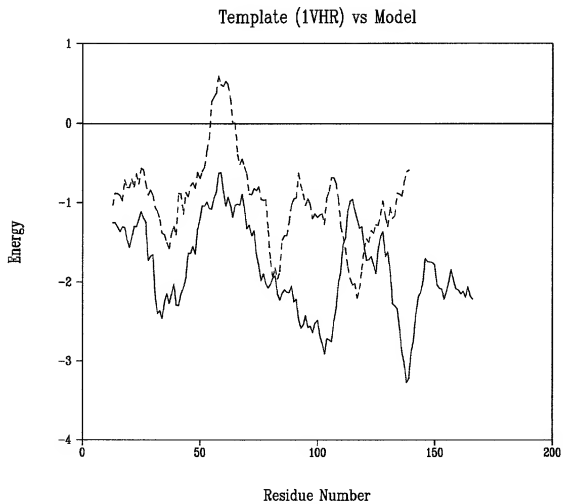
```

Figure 32



HPP2 Homology Model

Figure 33



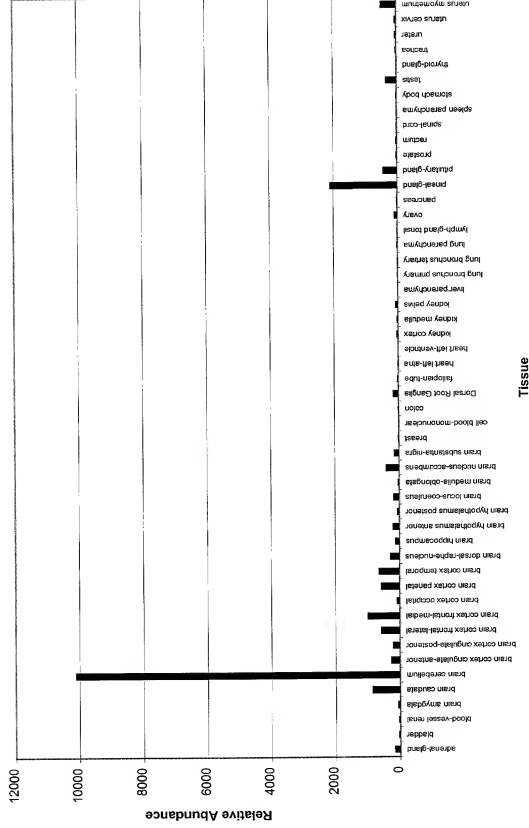


Figure 35

Relative Abundance of BMV_HPP5 in Normal Tissues

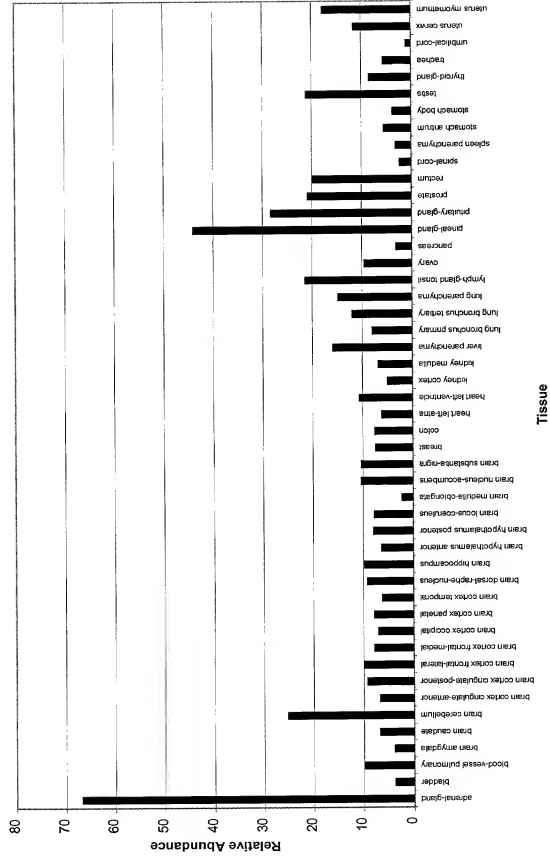


Figure 36

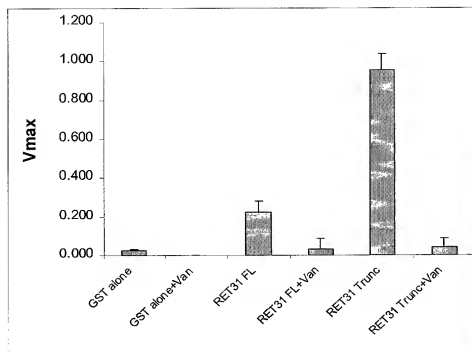


Figure 37

```

                                10      20      30
pdblmpk ASFFVEILPFLYLGCARDSTNLDVLEEFGIKYI
          : : : : : : : : : : : : : : : :
BMY_HPP5 SRCFPGLCCEGKSTLVPTCISQPCLEFVANIGFTRILPNLYLGCQRDVLNKELMQQNGIGYV
          130      140      150      160      170      180

                                40      50      60      70      80      90
pdblmpk LNVTPLNLPNLFENAGEFKYKQIPISDHSQNLQFFPEAISFIDEARGKNCGLVHSLAG
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
BMY_HPP5 LNASNTCPKP-DFIPESHFLRVPVNDSEFKILPWLDKSVDFIEKAKASNGCVLVHCLAG
          190      200      210      220      230      240

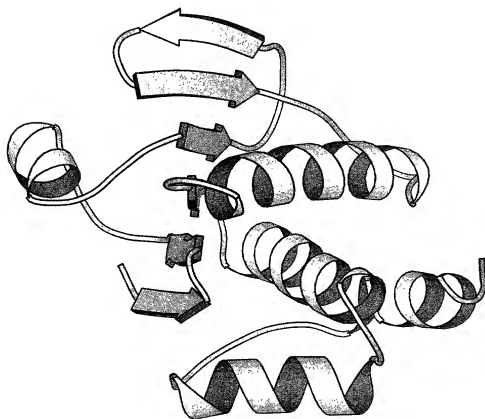
                                100      110      120      130      140
pdblmpk ISRSVITVTAYLMQKLNLSMNDAYDIVKMKKSNIISPNFNFMGQLLDFERTL
          : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
BMY_HPP5 ISRSATIAIAYIMKRMDSLD EAYRFVKEKRPTISPNFNFGLQLLAYEKKIKNQTGASGP
          250      260      270      280      290      300

BMY_HPP5 KSKLKLPLEKPNFVFAVSEGGQKSETPLSPPCADSATSEAAGQRPVHPASVFPVPSVQ
          310      320      330      340      350      360

```

100
 110
 120
 130
 140
 150
 160
 170
 180
 190
 200
 210
 220
 230
 240
 250
 260
 270
 280
 290
 300
 310
 320
 330
 340
 350
 360

Figure 38



HPP5 Homology Model

Figure 39

